FEUILLE D'EXERCICES N° 1. — GÉNÉRALITÉS

EXERCICE 1. — (i) Soient I un ensemble non vide d'indices et $(\mathcal{E}_i)_{i\in I}$ une famille de tribus sur un ensemble E. Montrer que

$$\bigcap_{i\in I}\mathcal{E}_i = \big\{A\in\mathcal{P}(E) : \text{pour tout } i\in I,\ A\in\mathcal{E}_i\big\} \qquad \text{est une tribu sur } E.$$

(ii) Soit $\mathcal{P}$ une partie de $\mathcal{P}(E)$. Montrer qu'il existe une unique tribu $\mathcal{E}$ sur E telle que :

a) $\mathcal{P}\subset\mathcal{E}$;

b) si $\mathcal{E}'$ est une tribu contenant $\mathcal{P}$, alors $\mathcal{E}\subset\mathcal{E}'$.

On note alors $\mathcal{E}=\sigma(\mathcal{P})$, c'est la *tribu engendrée* par $\mathcal{P}$.

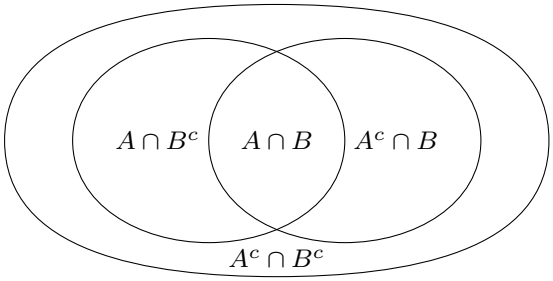
(iii) Si $A\in\mathcal{P}(E)\backslash\{\varnothing,E\}$, décrire $\sigma(A)$. Si A, $B\in\mathcal{P}(E)$, décrire $\sigma(A,B)$ en distinguant les différents cas pouvant se présenter. Déterminer le cardinal *maximal* d'une tribu engendrée par $n\geqslant 1$ parties quelconques de E.

Correction. — (i) Puisque pour tout $i\in I$, $\mathcal{E}_i$ est une tribu, $\varnothing\in E_i$. Ainsi, $\varnothing\in\bigcap_{i\in I}\mathcal{E}_i$ qui n'est pas vide. Soit $B\in\bigcap_{i\in I}\mathcal{E}_i$, alors pour tout $i\in I$, $B\in E_i$ et donc $B^c\in E_i$ puisque E_i est une tribu. Ainsi, $B^c\in\bigcap_{i\in I}\mathcal{E}_i$. Soient $B_1,\dots,B_n,\dots\in\bigcap_{i\in I}\mathcal{E}_i$, alors pour tout $i\in I$, $B_1,\dots,B_n,\dots\in E_i$ et donc $\bigcup_n B_n\in E_i$. Ainsi, $\bigcup_n B_n\in\bigcap_{i\in I}\mathcal{E}_i$. Ceci montre que $\bigcap_{i\in I}\mathcal{E}_i$ est une tribu sur E.

(ii) Considérons l'ensemble des tribus sur E qui contiennent $\mathcal{P}$ et notons le $(\mathcal{E}_i)_{i\in I}$. Cet ensemble — ou l'ensemble des indices I — est non vide puisque $\mathcal{P}(E)$ y appartient étant donné que c'est une tribu sur E qui contient $\mathcal{P}$. Soit $\mathcal{E}=\bigcap_{i\in I}\mathcal{E}_i$. C'est une tribu sur E, nous venons de le voir, et puisque $\mathcal{P}\subset\mathcal{E}_i$ pour tout $i\in I$, on a $\mathcal{P}\subset\mathcal{E}$. Donc $\mathcal{E}$ est une tribu qui contient $\mathcal{P}$. Si $\mathcal{E}'$ est une autre tribu sur E contenant P, il est alors clair que $\mathcal{E}\subset\mathcal{E}'$.

(iii) Si $A\subset E$ $A\neq\varnothing$ et $A\neq E$, on a $\sigma(A)=\{\varnothing,A,A^c,E\}$. En effet toute tribu contenant A contient ces 4 éléments, et ces 4 éléments forment une tribu à eux tous seuls.

Si A, $B\subset E$, on a $\sigma(A,B)=\{\varnothing,A,B,A^c,B^c,A\cap B,A\cap B^c,A^c\cap B,A^c\cap B^c,A\cup B,A\cup B^c,A^c\cup B,A^c\cup B^c,(A\cap B)\cup(A^c\cap B^c),(A\cap B^c)\cup(A^c\cap B),E\}$.



La liste est-t-elle correcte ? Pour comprendre ce qu'il se passe regardons les atomes de la tribu engendrée par A et B : ce sont les plus petits sous-ensembles non vides qu'on peut former avec A et B, soient $A\cap B$, $A\cap B^c$, $A^c\cap B$ et $A^c\cap B^c$. Ces atomes forment une partition de E et sont dans $\sigma(A,B)$, la tribu qu'ils engendrent est alors incluse dans $\sigma(A,B)$. Mais comme de plus $A=(A\cap B)\cup(A\cap B^c)$ et $B=(A\cap B)\cup(A^c\cap B)$, la tribu engendrée par les atomes contient $\sigma(A,B)$, ces deux tribus sont donc égales. La tribu engendrée par une partition finie (ou dénombrable) est facile à décrire : c'est l'ensemble des réunions (finies ou dénombrables)

des atomes ou éléments de la partition. Si la partition est de cardinal N (on ne compte pas les ensembles vides), la tribu engendrée est de cardinal 2^N, ce qui correspond au nombres de choix pour former des réunions. Pour $\sigma(A, B)$, la partition peut être de cardinal 1, 2, 3 ou 4 ; la tribu engendrée est suivant les cas de cardinal 2, 4, 8 ou 16.

Si on considère $A_1, \ldots, A_n \subset E$, les atomes s'écrivent $A^{\varepsilon_1} \cap \ldots \cap A^{\varepsilon_n}$, avec ε_i indiquant si on garde l'ensemble tel quel ou si on prend son complémentaire. Le nombre de telles intersections non vides est un entier k compris entre 1 et 2^n (pour tout $1 \leqslant k \leqslant 2^n$ [et $k \leqslant \operatorname{Card} \Omega$], on peut construire $A_1, \ldots A_n$ tels que le nombre d'intersections non vides soit précisément k). On a alors

$$\operatorname{Card} \sigma(A_1, \ldots, A_n) = 2^k \leqslant 2^{2^n}.$$

Notons au passage que ce cardinal croît très rapidement avec n. Ceci est peut-être à rapprocher du fait qu'il n'existe pas de tribu dénombrable...

Exercice 2. — Démontrer qu'il n'existe pas de tribu dénombrable.

Correction. — Soit $(E, \mathcal{E})$ un espace mesurable. Nous définissons la relation $\mathcal{R} : E \times E \to \{\text{vrai}, \text{faux}\}$ par

$$x \mathcal{R} y \iff \forall B \in \mathcal{E}, \quad (x \in B \wedge y \in B) \vee (x \notin B \wedge y \notin B)$$

ou en passant à la contraposée

$$x \not\mathcal{R} y \iff \exists B \in \mathcal{E}, \quad (x \in B \wedge y \notin B) \vee (x \notin B \wedge y \in B)$$

et puisque $\mathcal{E}$ est stable par passage au complémentaire, ceci se simplifie en

$$x \not\mathcal{R} y \iff \exists B \in \mathcal{E}, \quad x \in B \wedge y \notin B$$

En toutes lettres, $x \mathcal{R} y$ si et seulement si x et y ne peuvent pas être séparés par un ensemble mesurable. La relation $\mathcal{R}$ est une relation d'équivalence : elle est réflexive ($x \mathcal{R} x$ pour tout $x \in E$) ; elle est symétrique ($x \mathcal{R} y \iff y \mathcal{R} x$) ; et elle est transitive (si $x \mathcal{R} y$ et $y \mathcal{R} z$ alors $x \mathcal{R} z$, car si tout ensemble mesurable contenant x contient y et si tout ensemble mesurable contenant y contient z, alors tout ensemble mesurable contenant x contient z).

Les *atomes* de la tribu $\mathcal{E}$ sont les classes d'équivalence de la relation $\mathcal{R}$. Ils forment une partition de E. Explicitement, si $x \in E$, la classe d'équivalence de x est

$$\dot{x} = \bigcap_{B \in \mathcal{E},\, x \in B} B$$

car c'est l'ensemble des $y \in E$ tel que tout $B \in \mathcal{E}$ contenant x, contient y. On constate que sans hypothèse additionnelle, rien ne garantit la mesurabilité des atomes.

Supposons maintenant la tribu $\mathcal{E}$ au plus dénombrable. Il est alors clair que les atomes sont mesurables. Leur nombre est donc aussi au plus dénombrable et la tribu qu'ils engendrent est incluse dans $\mathcal{E}$. Pour montrer que ces deux tribus sont égales, montrons que si $B \in \mathcal{E}$, alors B s'écrit comme une réunion (au plus dénombrable) d'atomes.

Si $x \in B$, alors $\dot{x} \subset B$. Ainsi $B = \bigcup_{x \in B} \{x\} \supset \bigcup_{x \in B} \dot{x}$. L'inclusion réciproque étant évidente, on a donc égalité. Par hypothèse, cette réunion ne comporte qu'un nombre au plus dénombrable de termes distincts. Elle est donc, c'est-à-dire B, dans la tribu engendrée par les atomes. Nous avons donc égalité des deux tribus.

Pour éviter un choix apparent, soit $\dot{E}$ l'ensemble des classes d'équivalences qui est supposé au plus dénombrable. Ce que nous avons vu est que si $B \in \mathcal{E}$,

$$B = \bigcup_{\dot{x} \in \dot{E},\, \dot{x} \subset B} \dot{x}.$$

Pour finir, si $\dot{E}$ est fini, de cardinal n, nous savons que $\mathcal{E}$ est de cardinal 2^n et est donc finie. Si $\dot{E}$ est de cardinal infini dénombrable, $\aleph_0$, alors $\mathcal{E}$ est de cardinal $2^{\aleph_0}$ qui est égal à la

puissance du continu. Il n'existe pas de tribu de cardinal intermédiaire, donc, notamment, il n'existe pas de tribu dénombrable.

EXERCICE 3. — Soient Ω un ensemble, $(E,\mathcal{E})$ un espace mesurable et $X : \Omega \to E$ une application. Montrer que $\sigma(X) = \{\{X \in B\} : B \in \mathcal{E}\}$ est une tribu sur Ω et que c'est la plus petite tribu sur Ω tel que $X : \Omega \to E$ est mesurable, la tribu $\mathcal{E}$ sur E étant fixée.

Correction. — Puisque $\varnothing \in \mathcal{E}$, $X^{-1}(\varnothing = \varnothing$, on a $\varnothing \in \sigma(X)$. Si $A = X^{-1}(B) \in \sigma(X)$ avec $B \in \mathcal{E}$, alors $B^c \in \mathcal{E}$ et $X^{-1}(B^c) = X^{-1}(B)^c \in \sigma(X)$. Si $A_1 = X^{-1}(B_1)$, …, $A_n = X^{-1}(B_n)$, …, avec $B_n \in \mathcal{E}$, $\bigcup_n B_n \in \mathcal{E}$ et $X^{-1}(\bigcup_n B_n) = \bigcup_n X^{-1}(B_n) = \bigcup_n A_n \in \sigma(X)$ (il y a du choix dénombrable caché là dedans). Donc $\sigma(X)$ est une tribu sur Ω). Notons que X est $\mathcal{A}/\mathcal{E}$-mesurable si et seulement si $\sigma(X) \subset \mathcal{A}$. Comme X est $\sigma(X)/\mathcal{E}$-mesurable, $\sigma(X)$ est la plus petite $\mathcal{A}^X$ tribu sur Ω telle que $X : (\Omega, \mathcal{A}^X) \to (E,\mathcal{E})$ soit mesurable.

EXERCICE 4. — Soit $(E,\mathcal{E})$ un espace mesurable, F un ensemble et $f : E \to F$ une application. Montrer que $\tau(f) = \{C \subset F : f^{-1}(C) \in \mathcal{E}\}$ est une tribu sur $\mathcal{F}$, et que c'est la plus grande tribu sur F telle que f soit mesurable.

Correction. — Puisque $E = f^{-1}(F)$, on a $F \in \tau(f)$. Si $C \in \tau(f)$, puisque $f^{-1}(C^c) = f^{-1}(C)^c$, alors $C^c \in \tau(f)$. Enfin si $(C_n)_n \subset \tau(f)$, comme $f^{-1}(\bigcap_n C_n) = \bigcap_n f^{-1}(C_n)$, on a $\bigcap_n C_n \in \tau(f)$.

EXERCICE 5. — Soient $(\Omega,\mathcal{A})$ et $(E,\mathcal{E})$ deux espaces mesurables, une famille de parties $(B_i)_{i\in I} \subset \mathcal{E}$ telle que $\sigma((B_i)_{i\in I}) = \mathcal{E}$, et $X : \Omega \to E$ une application. Montrer que l'application X est $\mathcal{A}/\mathcal{E}$-mesurable si et seulement si pour tout $i \in I$, $X^{-1}(B_i) \in \mathcal{A}$.

Correction. — La condition est évidemment nécessaire. Elle est suffisante car par hypothèse la tribu $\tau(X)$ contient $(B_i)_{i\in I}$ et donc aussi $\sigma((B_i)_{i\in I}) = \mathcal{E}$.

EXERCICE 6. — Soient E et F deux espaces topologiques. Montrer que si $f : E \to F$ est continue alors f est borélienne, c'est-à-dire $\mathcal{E}/\mathcal{F}$-mesurable où $\mathcal{E} = \mathcal{B}(E)$ et $\mathcal{F} = \mathcal{B}(F)$ sont les tribus boréliennes de E et de F respectivement.

Correction. — Puisque f est continue, les images réciproques des ouverts de F par f sont des ouverts de E et sont donc dans $\mathcal{B}(E)$. La tribu $\tau(f)$ sur F contient donc tous les ouverts de F et donc $\mathcal{B}(F)$, ce qui montre que f est borélienne.

EXERCICE 7. — Soient I un ensemble non vide d'indices et $(E_i,\mathcal{E}_i)_{i\in I}$ une famille d'espaces mesurables non vides.

Le produit $E = \prod_{i\in I} E_i$ est muni de la tribu produit $\mathcal{E} = \bigotimes_{i\in I} \mathcal{E}_i$ engendrée par les pavés élémentaires $\prod_{i\in I} B_i$, $B_i \in \mathcal{E}_i$, et $B_i = E_i$ sauf pour un nombre fini d'indices $i \in I$.

Soient $(\Omega,\mathcal{A})$ un espace mesurable et $X_i : \Omega \to E_i$, $i \in I$ des applications.

Montrer que les $X_i : (\Omega,\mathcal{A}) \to (E_i,\mathcal{E}_i)$ sont mesurables si et seulement si $X = (X_i) : (\Omega,\mathcal{A}) \to (E,\mathcal{E})$ est mesurable.

Correction. — LEMME. — *Les applications coordonnées $f_i : E \to E_i$ sont $\mathcal{E}/\mathcal{E}_i$-mesurables.*

Démonstration du lemme. — Si $B_i \in \mathcal{E}_i$, $f_i^{-1}(B_i) = \prod_{j=1}^{i-1} E_j \times B_i \times \prod_{j=i+1}^{n} E_j$ qui est un pavé élémentaire et est donc dans $\mathcal{E}$. □

Démonstration. — Si $X : \Omega \to E$ est $\mathcal{A}/\mathcal{E}$-mesurable, d'après le lemme et par composition, $X_i = f_i \circ X : \Omega \to E_i$ est $\mathcal{A}/\mathcal{E}_i$-mesurable. Réciproquement, supposons les coordonnées X_i $\mathcal{A}/\mathcal{E}_i$-mesurables. La tribu $\tau(X)$ sur E contient tous les $\prod_{j=1}^{i-1} E_j \times B_i \times \prod_{j=i+1}^{n} E_j$ puisque $X^{-1}(\prod_{j=1}^{i-1} E_j \times B_i \times \prod_{j=i+1}^{n} E_j) = X_i^{-1}(B_i) \in \mathcal{A}$ par hypothèse. Donc $\mathcal{E} \subset \tau(X)$, ce qui montre que X est $\mathcal{A}/\mathcal{E}$-mesurable. □

Exercice 8 (lois binomiales négatives). — On considére le jeu classique et infini de « pile ou face » (schéma de Bernoulli infini), la probabilité d'obtenir pile étant $p \in]0,1[$.

(i) Décrire le modèle probabilisé filtré correspondant. On calculera aussi le cardinal de chaque tribu de la filtration associée.

(ii) Soit T_1 le premier instant d'obtention d'un « pile ». Montrer qu'il s'agit d'un temps d'arrêt et déterminer sa loi.

(iii) On définit T_n comme n-ième instant d'obtention de « pile ». Préciser sa définition, montrer qu'il s'agit d'un temps d'arrêt et déterminer sa loi.

(iv) Que dire si $p = 0$ ou $p = 1$?

Correction. — (i) Naturellement, l'ensemble fondamental minimal rendant compte des différentes possibilités est $\Omega = \{\text{pile}, \text{face}\}^{\mathbb{N}^*}$, ou encore $\Omega = \{0,1\}^{\mathbb{N}^*}$. Pour $n \geqslant 1$, $\mathcal{F}_n$ est la tribu engendrée par les n premières applications coordonnées sur Ω ; c'est une tribu finie dont les éléments s'écrivent $A \times \{0,1\}^{\mathbb{N}^*}$ avec $A \in \mathcal{P}(\{0,1\}^n)$; le cardinal de $\mathcal{F}_n$ est donc 2^n. La suite de tribus $\mathcal{F} = (\mathcal{F}_n)_{n\geqslant 1}$ est croissante, c'est donc une filtration sur Ω. La tribu ambiante $\mathcal{A}$ de Ω est $\mathcal{A} = \bigvee_{n\geqslant 1} \mathcal{F}_n$ qui est aussi la tribu engendrée par les applications coordonnées ou la tribu produit. La mesure de probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{A})$ est l'unique mesure de probabilité qui coïncide sur chaque $\mathcal{F}_n$ avec la mesure de probabilité associée au jeu de pile ou face répété n fois (Carathéodory). On codera 1 = pile et 0 = face.

(ii) Pour $\omega = (\omega_1, \ldots, \omega_n, \ldots) \in \Omega$, on pose $X_n(\omega) = \omega_n$, $n \geqslant 1$, X_n est la n-ième application coordonnée sur Ω. Soit $T_1 = \inf\{n \geqslant 1 : X_n = 1\}$. Pour tout $n \geqslant 1$, $\{T_1 \leqslant n\} = \bigcup_{1\leqslant k\leqslant n}\{X_k = 1\}$. Puisque pour tout $1 \leqslant k \leqslant n$, $\{X_k = 1\} \in \mathcal{F}_k \subset \mathcal{F}_n$, clairement $\{T_1 \leqslant n\} \in \mathcal{F}_n$ puisque c'est une réunion au plus dénombrable d'éléments de $\mathcal{F}_n$. Ainsi, T_1 est un temps d'arrêt dans la filtration $\mathcal{F}$. Pour $n \geqslant 1$, par indépendance des différents X_k, on a

$$\begin{aligned}\mathbb{P}\{T_1 = n\} &= \mathbb{P}\{X_1 = 0, \ldots, X_{n-1} = 0, X_n = 1\} \\ &= \mathbb{P}\{X_1 = 0\} \times \cdots \times \mathbb{P}\{X_{n-1} = 0\} \times \mathbb{P}\{X_n = 1\} = (1-p)^{n-1} \times p.\end{aligned}$$

La loi de T_1 est donc la loi géométrique de paramètre p.

(iii) Nous définissons par récurrence les temps d'arrêt $(T_n)_{n\geqslant 1}$. Le temps T_1 est défini. Pour $n \geqslant 1$, posons $T_{n+1} = \inf\{k \geqslant T_n + 1 : X_n = 1\}$. Pour $k \geqslant 1$, l'événement $\{T_n \leqslant k\}$ est réalisé si au moins n « pile » sont obtenus au cours des k premiers lancers. Donc,

$$\{T_n \leqslant k\} = \bigcup_{1\leqslant i_1 < \cdots < i_n \leqslant k} \{X_{i_1} = 1, \ldots, X_{i_n} = 1\}.$$

C'est une réunion finie (voire vide), donc au plus dénombrable, d'éléments de $\mathcal{F}_k$, donc, pour

tout $k \geqslant 1$, $\{T_n \leqslant k\} \in \mathcal{F}_k$, ce qui montre que T_n est un temps d'arrêt. Pour $k \geqslant 1$, on a

$$\begin{aligned}
\mathbb{P}\{T_n = k\} &= \mathbb{P}\left(\bigcup_{\substack{I \subset \{1,\dots,k-1\} \\ \text{Card } I = n-1}} \left(\bigcap_{i \in I} \{X_{i_{n-1}} = 1\} \right) \cap \left(\bigcap_{i \in \{1,\dots,k-1\} \setminus I} \{X_i = 0\} \right) \cap \{X_k = 1\} \right) \\
&= \sum_{\substack{I \subset \{1,\dots,k-1\} \\ \text{Card } I = n-1}} \mathbb{P}\left\{ \left(\bigcap_{i \in I} \{X_{i_{n-1}} = 1\} \right) \cap \left(\bigcap_{i \in \{1,\dots,k-1\} \setminus I} \{X_i = 0\} \right) \cap \{X_k = 1\} \right\} \\
&= \sum_{\substack{I \subset \{1,\dots,k-1\} \\ \text{Card } I = n-1}} \left(\prod_{i \in I} \mathbb{P}\{X_{i_{n-1}} = 1\} \right) \times \left(\prod_{i \in \{1,\dots,k-1\} \setminus I} \mathbb{P}\{X_i = 0\} \right) \times \mathbb{P}\{X_k = 1\} \\
&= \sum_{\substack{I \subset \{1,\dots,k-1\} \\ \text{Card } I = n-1}} p^{n-1} \times (1-p)^{k-n} \times p = \binom{k-1}{n-1} p^n (1-p)^{k-n}.
\end{aligned}$$

où on aura convenu qu'une somme vide est nulle, ce qui a lieu lorsque $k < n$. La loi de T_n est la loi binomiale négative de paramètres n et p. Remarquons que pour $n = 1$, on retrouve la loi géométrique de paramètre p.

(iv) Lorsque $p = 0$, on obtient presque sûrement toujours face, donc $T_1 = \infty$ presque sûrement et comme $T_n \geqslant T_1$ pour tout $n \geqslant 1$, on a $\mathbb{P}\{T_n = \infty, \text{ pour tout } n \geqslant 1\} = 1$. Lorsque $p = 1$, on obtient presque sûrement toujours pile, donc $T_1 = 1$ presque sûrement et $T_n = n$ presque sûrement pour tout $n \geqslant 1$.

Exercice 9. — Soient $X = (X_n)_{n \in \mathbb{N}}$ un processus adapté défini sur un espace probabilisé filtré $(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{F})$ et $T : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{N}$ un temps d'arrêt. Montrer que X_T est une variable aléatoire et qu'elle est de plus $\mathcal{F}_T$-mesurable. Étendre ce résultat lorsque le temps d'arrêt $T : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{N} \cup \{\infty\}$ n'est pas nécessairement fini.

Correction. — Soit $(E, \mathcal{E})$ l'espace d'états du processus X. Pour $B \in \mathcal{E}$, on a

$$X_T^{-1}(B) = \{X_T \in B\} = \bigcup_{k \geqslant 0} \{X_k \in B\} \cap \{T = k\}.$$

Puisque $\{X_k \in B\} \cap \{T = k\} \in \mathcal{F}_k \subset \mathcal{A}$, on constate que $\{X_T \in B\} \in \mathcal{A}$ pour tout $B \in \mathcal{E}$. Donc X_T est une variable aléatoire. De plus, pour $n \in \mathbb{N}$, on a

$$\{X_T \in B\} \cap \{T \leqslant n\} = \bigcup_{0 \leqslant k \leqslant n} \{X_k \in B\} \cap \{T = k\}$$

Puisque $\{X_k \in B\} \cap \{T = k\} \in \mathcal{F}_k \subset \mathcal{F}_n$, $0 \leqslant k \leqslant n$, on constate que $\{X_T \in B\} \cap \{T \leqslant n\} \in \mathcal{F}_n$ pour tout $n \geqslant 0$, et donc $\{X_T \in B\} \in \mathcal{F}_T$ pour tout $B \in \mathcal{E}$. Donc X_T est une variable aléatoire $\mathcal{F}_T$-mesurable.

Lorsque T n'est pas nécessairement fini, on ajoute à E un point à l'infini ou point cimetière : soient $\delta \notin E$, $\bar{E} = E \cup \{\delta\}, \bar{\mathcal{E}} = \sigma(\mathcal{E}, \{\delta\})$. On pose alors $X_\infty = \delta$. Soient $\bar{B} \in \mathcal{E}$ et $B = \bar{B} \cap E$. On a $\{X_T \in B\} \subset \{T < \infty\}$, et comme précédemment $\{X_T \in B\} \in \mathcal{F}_T \subset \mathcal{A}$. De plus $\{X_T = \delta\} = \{T = \infty\} \in \mathcal{F}_T \subset \mathcal{A}$. Donc, que δ appartiennent ou non à $\bar{B}$, $\{X_T \in \bar{B}\} \in \mathcal{F}_T \subset \mathcal{A}$.

Exercice 10. — Soient $(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{F})$ un espace probabilisé filtré avec $\mathcal{F} = (\mathcal{F}_n)_{n \geqslant 0}$, et S, T deux temps d'arrêt, alors $S + T$ est un temps d'arrêt. Cette propriété vous paraît-elle vraie si l'ensemble des temps est $\mathbb{Z}, \dots, \mathbb{R}_+$?

Correction. — Pour $n \in \mathbb{N}$, on a

$$\{S+T \leqslant n\} = \bigcup_{k=0}^{n} \{S+k \leqslant n, T=k\} = \bigcup_{k=0}^{n} \{S \leqslant n-k\} \cap \{T=k\}.$$

Or, $\{S \leqslant n-k\} \in \mathcal{F}_{n-k} \subset \mathcal{F}_n$ et $\{T=k\} \in \mathcal{F}_k \subset \mathcal{F}_n$, donc $\{S+T \leqslant n\} \in \mathcal{F}_n$. La variable aléatoire $S+T$ est un temps d'arrêt.

Remarque. — Dans ce qui précède, il suffit que T soit un temps d'arrêt positif n'ayant qu'un nombre au plus dénombrable de valeurs possibles, l'autre temps d'arrêt étant quelconque mais *positif*.

Si l'ensemble des temps est $\mathbb{Z}$, la somme de deux temps d'arrêt n'est pas nécessairement un temps d'arrêt. Prenons par exemple pour T le temps d'arrêt constant égal à -1. On a alors $\{S+T \leqslant n\} = \{S \leqslant n+1\} \in \mathcal{F}_{n+1}$ mais n'est pas nécessairement $\mathcal{F}_n$-mesurable.

Si l'ensemble des temps est $\mathbb{R}_+$, on peut être tenté de raisonner comme dans le cas de $\mathbb{N}$. Alors on obtient des réunions éventuellement strictement plus que dénombrables dont on ne peut rien dire quant à la mesurabilité. Supposons ne pas être dans le cadre de la remarque précédente. On peut essayer de discrétiser l'un de ces temps d'arrêt : soient $\varepsilon > 0$ et $T_\varepsilon = \varepsilon\lfloor T/\varepsilon \rfloor$, autrement dit

$$T_\varepsilon = n\varepsilon \iff T \in [n\varepsilon, (n+1)\varepsilon[,$$

notamment, $T_\varepsilon \leqslant T$. Ainsi, $\{S+T \leqslant t\} \subset \{S+T_\varepsilon \leqslant t\}$, et de plus $\{S+T_\varepsilon \leqslant t\}$ décroît vers $\{S+T \leqslant t\}$ lorsque ε tend vers 0. Pour $t \geqslant 0$, on a

$$\{S+T_\varepsilon \leqslant t\} = \bigcup_{\{n : 0 \leqslant n\varepsilon \leqslant t\}} \{S+n\varepsilon \leqslant t, T_\varepsilon = n\varepsilon\} = \bigcup_{\{n : 0 \leqslant n\varepsilon \leqslant t\}} \{S \leqslant t-n\varepsilon\} \cap \{T_\varepsilon = n\varepsilon\}.$$

Pour n tel que $0 \leqslant n\varepsilon \leqslant t$, on a $\{S \leqslant t-n\varepsilon\} \in \mathcal{F}_{t-n\varepsilon} \subset \mathcal{F}_t \subset \mathcal{F}_{t+\varepsilon}$ et $\{T_\varepsilon = n\varepsilon\} = \{T \in [n\varepsilon, (n+1)\varepsilon[\in \mathcal{F}_{(n+1)\varepsilon} \subset \mathcal{F}_{t+\varepsilon}$. Alors,

$$\{S+T \leqslant t\} = \bigcap_{\varepsilon \in \mathbb{Q}_+^*} \{S+T_\varepsilon \leqslant t\} \in \mathcal{F}_{t+} = \bigcap_{\varepsilon \in \mathbb{Q}_+^*} \mathcal{F}_{t+\varepsilon}.$$

Si l'intersection de ces tribus est $\mathcal{F}_t$ pour tout t, on dit alors que la filtration est continue à droite, la variable aléatoire $S+T$ est alors un temps d'arrêt de la filtration.

Remarque. — Toute filtration $\mathcal{F}$ peut être rendue continue à droite en considérant $\mathcal{F}_+ = (\mathcal{F}_{t+})_{t \geqslant 0}$. On a bien entendu $(\mathcal{F}_+)_+ = \mathcal{F}_+$, ou $\mathcal{F}_+ = \mathcal{F}$ si $\mathcal{F}$ est continue à droite.

Exercice 11 *(identité de Wald).* — On considère sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ une suite de variables aléatoires réelles, indépendantes $X_n : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $n \in \mathbb{N}^*$, intégrables de même espérance $\mathbb{E}[X_n] = m$. La filtration considérée est celle du processus X, avec $\mathcal{F}_0 = \{\varnothing, \Omega\}$ la tribu triviale. Dans cette filtration, nous considérons un temps d'arrêt $T : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{N}^*$ qui est supposé d'espérance finie ($\mathbb{E}[T] < \infty$).

Montrer que

$$\mathbb{E}\left[\sum_{n=1}^{T} X_n\right] = \mathbb{E}[T] \times m.$$

(On pourra commencer par considérer les X_n de signe constant et égal, puis passer au cas général.)

Correction. — Supposons $X_n \geqslant 0$ pour tout $n \geqslant 1$ afin de ne pas se poser de problèmes d'intégrabilité. On a

$$\mathbb{E}\left[\sum_{n=1}^{T} X_n\right] = \mathbb{E}\left[\sum_{n=1}^{\infty} \mathbb{1}_{\{T \leqslant n\}} X_n\right] = \sum_{n=1}^{\infty} \mathbb{E}\left[\mathbb{1}_{\{T \leqslant n\}} X_n\right],$$

par le théorème de Fubini–Tonelli. Or, pour tout $n \geqslant 1$,

$$\{n \leqslant T\} = \{T \geqslant n\} = \{T < n\}^c = \{T \leqslant n-1\}^c \in \mathcal{F}_{n-1}$$

(noter que ceci est vrai même pour $n = 1$). Comme $\mathcal{F}_{n-1}$ et X_n sont indépendantes, il en est de même de $\mathbb{1}_{\{T\leqslant n\}}$ et de X_n. Ainsi, comme conséquence du théorème de Fubini–Tonelli,

$$\mathbb{E}\big[\mathbb{1}_{\{T\leqslant n\}} X_n\big] = \mathbb{E}\big[\mathbb{1}_{\{T\leqslant n\}}\big] \times \mathbb{E}[X_n] = \mathbb{P}\{T \geqslant n\} \times m.$$

D'où,

$$\mathbb{E}\bigg[\sum_{n=1}^{T} X_n\bigg] = \sum_{n=1}^{\infty} \mathbb{P}\{T \geqslant n\} \times m = \mathbb{E}[T] \times m.$$

Pour passer au cas où les $(X_n)_{n\geqslant 1}$ sont intégrables non nécessairement de signe constant, on applique le résultat précédent à la suite des valeurs absolues pour constater que toutes les quantités calculées sont finies, puis on enlève les valeurs absolues en justifiant quelques manipulations de calcul par le théorème de Fubini.

Remarque. — On rappelle que si $T : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{N}$ est une variable aléatoire entière,

$$\sum_{n\geqslant 1} \mathbb{P}\{T \geqslant n\} = \sum_{n\geqslant 1} \sum_{k\geqslant n} \mathbb{P}\{T = k\} = \sum_{k\geqslant 1} \sum_{n=1}^{k} \mathbb{P}\{T = k\} = \sum_{k\geqslant 1} k\, \mathbb{P}\{T = k\} = \mathbb{E}[T],$$

l'interversion des sommes est possible par le théorème de Fubini–Tonelli. La seule difficulté est de convenablement indexer la nouvelle double somme obtenue.

Exercice 12. — Soient $X = (X_n)_{n\geqslant 0}$ un processus défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans un espace d'états $(E, \mathcal{E})$ et $A \in \mathcal{E}$ un sous-ensemble mesurable de E.

(i) Le temps atteinte de A par le processus X est

$$T^A = \inf\{n \geqslant 0 : X_n \in A\}.$$

On définit naturellement T_r^A le temps de r-ième atteinte de A pour $r \geqslant 1$. Montrer que ce sont des temps d'arrêt pour la filtration naturelle $\mathcal{F}^X$ du processus X.

(ii) Le temps de dernier retour dans A par le processus X est

$$L^A = \sup\{n \geqslant 0 : X_n \in A\}.$$

Constater que rien n'indique que ce soit un temps d'arrêt en général pour $\mathcal{F}^X$.

(iii) Supposons le processus déterministe : $X = (x_n)_{n\geqslant 0}$ est une suite de variables aléatoires constantes. Quelle est sa filtration naturelle ? Que dire des temps d'atteinte et de dernier retour ?

(iv) On suppose maintenant que le processus X est à accroissements indépendants à valeurs dans $\mathbb{Z}$, et que plus précisément $X_n = X_0 + \xi_1 + \cdots + \xi_n$ où X_0 et les $(\xi_n)_{n\geqslant 1}$ sont des variables aléatoires indépendantes à valeurs dans $\mathbb{Z}$, et $(\xi_n)_{n\geqslant 1}$ identiquement distribuées de loi $(1-p)\delta_{\{-1\}} + p\delta_{\{1\}}$ ($X = (X_n)_{n\geqslant 0}$ est une marche aléatoire simple [non nécessairement symétrique] sur $\mathbb{Z}$) avec $p \in\,]0, 1[$. Discuter, pour $A = \{\ldots, -2, -1\} = \mathbb{Z}_-^*$, du fait que L^A puisse être un temps d'arrêt du processus X ou non.

Correction. — Soit $X_n : (\Omega, \mathcal{A}, \mathbb{P}) \to (E, \mathcal{E})$, $n \geqslant 0$, un processus de filtration naturelle $\mathcal{F}^X$ et soit $A \in \mathcal{E}$.

(i) Soient $T^A = \inf\{n \geqslant O : X_n \in A\}$, avec $\inf \varnothing = \infty$, et par récurrence, $T_1^A = T^A$, $T_{r+1}^A = \inf\{n \geqslant T_r^A + 1 : X_n \in A\}$, $r \geqslant 1$. Pour $r \geqslant 1$ et $n \in \mathbb{N}$, on a

$$\begin{aligned}\big\{T_r^A \leqslant n\big\} &= \{\exists\, 0 \leqslant n_1 < \cdots < n_r \leqslant n : X_{n_1} \in A, \ldots, X_{n_r} \in A\} \\ &= \bigcup_{0\leqslant n_1<\cdots<n_r\leqslant n} \{X_{n_1} \in A\} \cap \ldots \cap \{X_{n_r} \in A\}.\end{aligned}$$

Comme tous les n_i sont inférieurs ou égaux à n, chacun des événements $\{X_{n_i} \in A\}$ appartient à $\mathcal{F}_n$, la réunion de leurs intersections écrite plus haut aussi puisqu'elle ne porte que sur un nombre fini et donc au plus dénombrable d'indices. Puisque c'est vrai pour tout $n \in \mathbb{N}$, T_r^A est un $\mathcal{F}^X$ temps d'arrêt pour tout $r \geqslant 1$ (c'est notamment une variable aléatoire).

(ii) Soit $L^A = \sup\{n \leqslant 0 : X_n \in A\}$, où on convient que $\sup \varnothing = 0-$ (ou encore $-\infty$, -1,...), convention à comparer avec $\inf \varnothing = \infty$, l'ensemble des temps étant $\mathbb{N}$). Pour $n \in \mathbb{N}$,

$$\big\{L^A = n\big\} = \{X_n \in A\} \cap \bigcap_{k>n} \{X_k \notin A\} \in \mathcal{A},$$

notamment L^A est bien une variable aléatoire (discrète). Mais $\{L^A \leqslant n\}$ semble *a priori* dépendre du futur strict à l'instant n.

Exemple. — Soient $(\xi_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées de loi $(\delta_{\{-1\}} + \delta_{\{1\}})/2$ de filtration naturelle $\mathcal{F}$ avec $\mathcal{F}_0 = \{\varnothing, \Omega\}$. Posons $Y_0 = 0$, et pour $n \geqslant 1$, $Y_n = \xi_1 + \cdots + \xi_n$. Le processus Y est la marche aléatoire symétrique sur $\mathbb{Z}$ issue de 0. Posons $T = \inf\{n \geqslant 0 : Y_n \in \{-2, 2\}$ qui est un temps d'arrêt presque sûrement fini, et $X_n = Y_{n\wedge T}$, $n \geqslant 0$, autrement dit le processus X est le processus Y arrêté en T, ou même encore absorbé par $A = \{-2, 2\}$ puisqu'une fois A atteint, le processus X y demeure. Considérons $L^{\{0\}}$ le dernier temps de retour de X en 0. C'est une variable aléatoire presque sûrement finie et on a

$$\big\{L^{\{0\}} = n\big\} = \{X_n = 0\} \cap \big(\{\xi_{n+1} = 1, \xi_{n+2} = 1\} \cup \{\xi_{n+1} = -1, \xi_{n+2} = -1\}\big).$$

En particulier

$$\big\{L^{\{0\}} = 0\big\} = \{\xi_1 = 1, \xi_2 = 1\} \cup \{\xi_1 = -1, \xi_2 = -1\} \notin \mathcal{F}_0 = \{\varnothing, \Omega\}$$

puisque cet événement est de probabilité 1/2. Le dernier temps de retour en 0 dans cette situation n'est donc pas un temps d'arrêt.

Remarque. — Si un dernier temps de retour dépend du présent et du futur, il ne dépend pas du passé strict. Ainsi si on retourne le sens du temps, on obtient alors un temps d'atteinte qui est un temps d'arrêt de la nouvelle filtration obtenue par retournement du temps.

(iii) Le processus $X = (X_n)_{n\geqslant 0}$ est supposé déterministe, c'est-à-dire qu'il existe une suite $(x_n)_{n\geqslant 0}$ d'éléments de E tels qu'on a identiquement $X_n = x_n$ pour tout $n \geqslant 0$.

La filtration naturelle de X est alors la filtration constante : $\mathcal{F}_n^X = \{\varnothing, \Omega\}$ pour tout $n \geqslant 0$. Ainsi si T est un $\mathcal{F}^X$-temps d'arrêt, pour tout $n \in \mathbb{N}$, $\{T = n\} \in \{\varnothing, \Omega\}$. Il n'existe alors qu'un seul entier $t \geqslant 0$ tel que $\{T = t\} = \Omega$, et donc T est constant, égal à t.

Remarque. — Ce type de filtration est assez peu intéressante puisque $\mathcal{F}_\infty = \{\varnothing, \Omega\}$ et toute variable aléatoire $\mathcal{F}_\infty$-mesurable est constante.

(iv) Soient $X_n = X_0 + \xi_1 + \cdots + \xi_n$ comme défini dans l'énoncé et $A = \mathbb{Z}_-^*$.

Si $p \neq 1/2$, la loi forte des grands nombres montre que X_n/n tend presque surement vers $\mathbb{E}[\xi_1] = 2p - 1 \neq 0$, donc $X_n \sim (2p-1)n$ quand n tend vers l'infini.

Lorsque $p < 1/2$, X_n tend vers $-\infty$ presque sûrement, notamment le processus X demeure presque sûrement dans $A = \mathbb{Z}_-^*$ à partir d'un certain rang aléatoire. Ainsi $L^A = \infty$ presque sûrement. En faisant fi des négligeables (ou en complétant la filtration), on constate que L^A est un temps d'arrêt.

Lorsque $p = 1/2$, le comportement asymptotique de X_n n'est plus décrit par la loi forte des grands nombres. Nous utilisons un résultat classique sur la marche aléatoire symétrique sur $\mathbb{Z}$:

$$\liminf_{n\to\infty} X_n = -\infty \quad \text{p.s.} \qquad \text{et} \qquad \limsup_{n\to\infty} X_n = +\infty \quad \text{p.s.}$$

On a donc $L^A = \infty$ presque sûrement et une conclusion semblable au cas $p < 1/2$.

Lorsque $1/2 < p < 1$, X_n tend vers $+\infty$ presque sûrement, et L^A est fini presque sûrement. Soit $n \geqslant 0$ tel que $\mathbb{P}\{L^A = n\} > 0$. On a

$$\{L^A = n\} = \{X_n = 0\} \cap \bigcap_{k \geqslant 1} \{\xi_{n+1} + \cdots + \xi_{n+k} > 0\}$$

et donc, par indépendance,

$$0 < \mathbb{P}\{L^A = n\} = \mathbb{P}\{X_n = 0\} \times \mathbb{P}\Big(\bigcap_{k \geqslant 1} \{\xi_{n+1} + \cdots + \xi_{n+k} > 0\}\Big).$$

Or, pour tout $m \geqslant 0$,

$$\mathbb{P}\Big(\bigcap_{k \geqslant 1} \{\xi_{n+1} + \cdots + \xi_{n+k} > 0\}\Big) = \mathbb{P}\Big(\bigcap_{k \geqslant 1} \{\xi_1 + \cdots + \xi_k > 0\}\Big)$$

nombre strictement positif que nous notons α. Remarquons que $\alpha \in]0, 1[$ puisque

$$\alpha = 1 - \mathbb{P}\{\exists\, k \geqslant 1 : \xi_1 + \cdots + \xi_k \leqslant 0\} \leqslant 1 - \mathbb{P}\{\xi_1 = -1\} = p < 1.$$

Or, $\mathbb{P}^0\{L^A = 0\} = \alpha$. Ainsi, $\{L^A = 0\}$ n'est pas $\mathcal{F}_0^X$-mesurable puisque tous les événements $\mathcal{F}_0^X$-mesurables (ils sont de la forme $\{X_0 \in A,\, A \subset \mathbb{N}$) sont de probabilité 0 ou 1 sous $\mathbb{P}^0$.

Pour finir, si $p = 1$, alors $L^A = -X_0$ si $X_0 \leqslant 0$ et $L^A = \infty$ si $X_0 > 0$ presque sûrement. Donc L^A est $\mathcal{F}_0$-mesurable (aux négligeables près), c'est donc un temps d'arrêt.

Remarque. — Notons que, puisque $\{L^A = n\} \subset \{X_n = 0\}$, la relation $\mathbb{P}\{L^A = n\} = \mathbb{P}\{X_n = 0\} \times \alpha$ s'écrit $\mathbb{P}\{L^A = n \,|\, X_n = 0\} = \alpha = \mathbb{P}^0\{L^A = 0\}$ qui est une relation typiquement markovienne.

Exercice 13. — Le processus X considéré correspond encore au schéma de Bernoulli infini avec un paramètre $p \in]0, 1[$.

(i) On définit

$$T = \begin{cases} 2 & \text{si } X_0 = 1, \\ 3 & \text{si } X_0 = 0. \end{cases}$$

Montrer que c'est un temps d'arrêt dans la filtration naturelle de X.

(ii) On définit

$$T' = \begin{cases} 2 & \text{si } X_3 = 1, \\ 3 & \text{si } X_3 = 0. \end{cases}$$

Montrer que ce n'est pas un temps d'arrêt dans cette même filtration.

(iii) Que dire de T' si $p = 0$ ou 1 ? (On pourra supposer les variables aléatoires identiquement égales à leurs valeurs presque sûres.)

Correction. — (i) On a $\{T = 2\} = \{X_0 = 1\} \in \mathcal{F}_0 \subset \mathcal{F}_2$, $\{T = 3\} = \{X_0 = 0\} \in \mathcal{F}_0 \subset \mathcal{F}_3$ et $\{T = n\} = \varnothing \in \mathcal{F}_n$ pour tout $n \in \mathbb{N}\backslash\{2, 3\}$. Donc T est un temps d'arrêt.

(ii) On a $\{T' = 2\} = \{X_3 = 1\} \notin \mathcal{F}_2$. En effet, si $\{X_3 = 1\} \in \mathcal{F}_2$, alors $\{X_3 = 1\}$ serait indépendant de lui-même puisque $\mathcal{F}_2 = \sigma\{X_1, X_2\}$ et X_3 sont indépendants, on aurait alors $\mathbb{P}\{X_3 = 1\} = p \in \{0, 1\}$, ce qui n'est pas le cas par hypothèse. En revanche, $\{T' = 3\} = \{X_3 = 0\} \in \mathcal{F}_3$ et $\{T' = n\} = \varnothing \in \mathcal{F}_n$ pour tout $n \in \mathbb{N}\backslash\{2, 3\}$. Cependant, T' n'est pas un temps d'arrêt.

(iii) Pour $p = 0$ ou $p = 1$ la filtration (complétée) est constante. Donc toute variable est (presque sûrement) constante et $\mathcal{F}_0$-mesurable. C'est le cas de T' qui est alors un temps d'arrêt.

Exercice 14. — Considérons une marche aléatoire $X = (X_n)_{n \geqslant 0}$ simple sur $\mathbb{Z}$, symétrique ($p = 1/2$) et issue de 0. Soit T le premier temps d'atteinte de $A = \{1\}$ par la marche X.

(i) Établir que

$$\mathbb{P}\big\{T \leqslant 2k+1, X_{2k+1} \leqslant 1\big\} = \mathbb{P}\big\{T \leqslant 2k+1, X_{2k+1} \geqslant 1\big\}. \qquad (*)$$

(ii) Montrer que

$$\mathbb{P}\{T \leqslant 2k+1\} = 1 - \mathbb{P}\{X_{2k+1} = -1\}. \qquad (**)$$

(iii) En déduire la fonction de répartition de T ainsi que sa loi.

Correction. — Soient $(\xi_k)_{k\geqslant 1}$ une suite de variables aléatoire réelles indépendantes identiquement distribuées de loi $(\delta_{\{1\}} + \delta_{\{1\}})/2$, $X_0 = 0$ et $X_n = \xi + \cdots + \xi_n$ pour tout $n \geqslant 1$, autrement dit $X = (X_n)_{n\geqslant 0}$ est une marche aléatoire simple et symétrique sur $\mathbb{Z}$.

Soient $A = \{1\}$ et $T = T^{\{1\}} = \inf\{n \geqslant 0 : X_n = 1\}$ qui est un temps d'arrêt (fini comme on le verra) et qui ne prend que des valeurs impaires.

(i) On a pour $i \leqslant k$

$$\{T = 2i+1, X_{2k+1} \leqslant 1\} = \{T = 2i+1\} \cap \{\xi_{2i+2} + \cdots + \xi_{2k+1} \leqslant 0\}$$

qui est l'intersection de deux événements indépendants, donc

$$\begin{aligned}\mathbb{P}\{T = 2i+1, X_{2k+1} \leqslant 1\} &= \mathbb{P}\{T = 2i+1\} \times \mathbb{P}\{\xi_{2i+2} + \cdots + \xi_{2k+1} \leqslant 0\} \\ &= \mathbb{P}\{T = 2i+1\} \times \mathbb{P}\{\xi_{2i+2} + \cdots + \xi_{2k+1} \leqslant 0\} = \mathbb{P}\{T = 2i+1, X_{2k+1} \geqslant 1\}\end{aligned}$$

où les dernières inégalités découlent de la symétrie de la marche. En sommant sur i de 0 à k, on obtient

$$\mathbb{P}\{T \leqslant 2k+1, X_{2k+1} \leqslant 1\} = \mathbb{P}\{T \leqslant 2k+1, X_{2k+1} \geqslant 1\}. \qquad (*)$$

(ii) En utilisant $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$,

$$\begin{aligned}\mathbb{P}\{T \leqslant 2k+1\} &= \mathbb{P}\{T \leqslant 2k+1, X_{2k+1} \leqslant 1\} \\ &\quad + \mathbb{P}\{T \leqslant 2k+1, X_{2k+1} \geqslant 1\} - \mathbb{P}\{T \leqslant 2k+1, X_{2k+1} = 1\} \\ &= 2\,\mathbb{P}\{T \leqslant 2k+1, X_{2k+1} > 1\} + \mathbb{P}\{T \leqslant 2k+1, X_{2k+1} = 1\}.\end{aligned}$$

Or, si $X_{2k+1} > 1$, alors $T \leqslant 2k+1$, et si $X_{2k+1} = 1$ alors $T \leqslant 2k+1$, donc

$$\{T \leqslant 2k+1, X_{2k+1} > 1\} = \{X_{2k+1} > 1\} \quad \text{et} \quad \{T \leqslant 2k+1, X_{2k+1} = 1\} = \{X_{2k+1} = 1\},$$

et ainsi

$$\begin{aligned}\mathbb{P}\{T \leqslant 2k+1\} &= 2\,\mathbb{P}\{X_{2k+1} > 1\} + \mathbb{P}\{X_{2k+1} = 1\} \\ &= \mathbb{P}\{X_{2k+1} > 1\} + \mathbb{P}\{X_{2k+1} < -1\} + \mathbb{P}\{X_{2k+1} = 1\} \\ &= 1 - \mathbb{P}\{X_{2k+1} = -1\}\end{aligned}$$

où on a utilisé la symétrie de la marche par rapport à 0 et le fait que X_{2k+1} ne prend que des valeurs impaires.

(iii) Or $\mathbb{P}\{X_{2k+1} = -1\}$ s'obtient par simple comptage : on part de 0 et en $2k+1$ pas on atteint -1 ; on effectue donc k pas positifs et $k+1$ pas négatifs ; il y a C_{2k+1}^k manière de ce faire chacune de probabilité $1/2^{2k+1}$; et donc $\mathbb{P}\{X_{2k+1} = -1\} = C_{2k+1}^k/2^{2k+1}$. Finalement, la fonction de répartition F_T est constante sur les intervalles $[2k+1, 2k+3[$ et on a, pour tout $k \geqslant 0$,

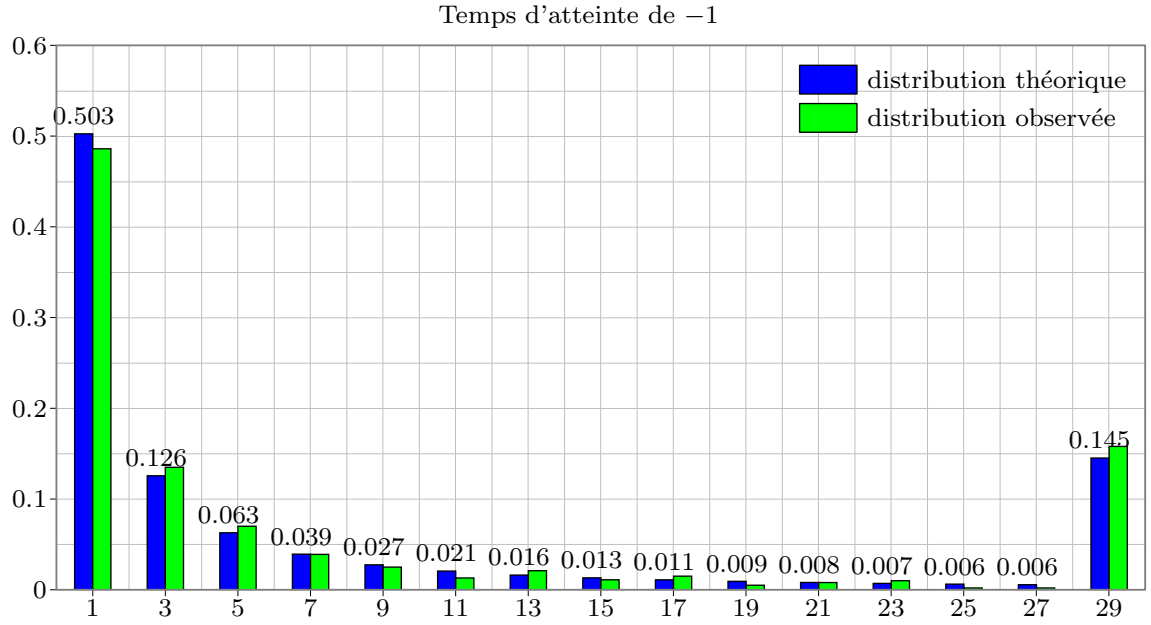
$$F_T(2k+1) = \mathbb{P}\{T \leqslant 2k+1\} = 1 - \frac{C_{2k+1}^k}{2^{2k+1}}\,.$$

On peut constater au passage que F_T tend vers 1 en $+\infty$, ce qui confirme que T est fini presque sûrement.

On a $\mathbb{P}\{T = 1\} = 1/2$, et pour $k \geqslant 1$,

$$\mathbb{P}\{T = 2k+1\} = \mathbb{P}\{T \leqslant 2k+1\} - \mathbb{P}\{T \leqslant 2k-1\} = \frac{C_{2k-1}^{k-1}}{2^{2k-1}} - \frac{C_{2k+1}^k}{2^{2k+1}}\,.$$

Pour finir, quelques simulations. Nous avons simulé 1000 trajectoires de la marche aléatoire simple symétrique arrêtée au temps d'atteinte de 1, pour obtenir un échantillon de taille 1000 de la variable T pour lequel nous avons considéré tous les cas du type $T \geqslant t_{\max}$ comme équivalents sur le graphique à $T = t_{\max}$. Les résultats sont les suivants :



En bleu, nous avons la distribution théorique. En notant que $\mathbb{P}\{T = 1\} = 1/2$,

$$\mathbb{P}\{T = 2k+1\} = \mathbb{P}\{T \leqslant 2k+1\} - \mathbb{P}\{T \leqslant 2k-1\} = \frac{C_{2k-1}^{k-1}}{2^{2k-1}} - \frac{C_{2k+1}^{k}}{2^{2k+1}} = \frac{2k+1}{2k+2}\,\mathbb{P}\{T = 2k-1\}$$

et

$$\mathbb{P}\{T \geqslant 2M+1\} = 1 - \mathbb{P}\{T < 2M+1\} = 1 - \mathbb{P}\{T \leqslant 2M-1\} = \frac{C_{2M-1}^{M-1}}{2^{2M-1}}\,,$$

(sur le graphique $t_{\max} = 2M+1 = 29$) sa détermination numérique se fait correctement de manière récurrente.

Exercice 15 (le collectionneur de vignettes). — Pierre achète des tablettes de chocolat d'une certaine marque. Dans chaque tablettes, il trouve une vignette qu'il peut coller dans un album édité par cette même marque. L'album contient m emplacements. Soient $T_1 = 1$, puis T_2 le rang d'obtention d'une vignette différente de la première, puis T_3 celui d'obtention d'une vignette différente des deux première, etc., jusqu'à T_m le rang de complétion de l'album. On pose $U_1 = T_1 = 1$, $U_2 = T_2 - T_1, \ldots, U_m = T_m - T_{m-1}$.

(i) Modéliser le processus d'obtention des vignettes avec des hypothèses raisonnables.

(ii) Déterminer la loi de la variable aléatoire discrète $(U_1, \ldots, U_n)$.

(iii) En déduire que les variables $U_1, \ldots, U_n$ sont indépendantes.

(iv) En déduire la fonction génératrice de T_m ainsi que son espérance.

Correction. — *Version rapide.* — Il est clair que les $U_1, \ldots, U_m$ sont (indépendantes) de loi respectives $\mathcal{G}((m-k+1)/m)$. On a donc

$$\begin{aligned}\mathbb{E}[T] &= \mathbb{E}[U_1 + \cdots + U_m] = \mathbb{E}[U_1] + \cdots + \mathbb{E}[U_m] \\ &= 1 + \frac{m-1}{m} + \frac{m-2}{m} + \cdots + \frac{1}{m} = m\Big(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m}\Big)\,.\end{aligned}$$

Version laborieuse. — (i) Nous considérons $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes et uniformément distribuées sur $\{1, \ldots, m\}$ (X_n désigne le numéro de la n-ième image

obtenue). Les variables $T_1, \dots, T_m$ et $U_1, \dots U_m$ sont celles de l'énoncé : $T_1 = 1$, $T_k = \inf\{n > T_{k-1} : X_n \notin \{X_{T_1}, \dots, X_{T_{k-1}}\}$ pour $2 \leqslant k \leqslant m$; $U_1 = 1$, $U_k = T_k - T_{k-1}$ pour $2 \leqslant k \leqslant m$.

(ii) Déterminons la loi conjointe des variables aléatoires discrètes $U_1, \dots, U_m$ (à valeurs dans $\mathbb{N}^*$) : soient $u_1, \dots, u_m \in \mathbb{N}^*$, alors $U_1 = u_1, \dots, U_m = u_m$ équivaut à l'existence d'une permutation σ de $\{1, \dots, m\}$ telle que

$$\begin{aligned}
&X_{u_1} = \sigma(1), X_{u_1+1} = \sigma(1), \dots,\\
&X_{u_1+u_2-1} = \sigma(1), X_{u_1+u_2} = \sigma(2), X_{u_1+u_2+1} \in \{\sigma(1), \sigma(2)\}, \dots,\\
&X_{u_1+u_2+u_3-1} \in \{\sigma(1), \sigma(2)\}, X_{u_1+u_2+u_3} = \sigma(3), \dots\\
&\dots, X_{u_1+\dots+u_m-1} \in \{\sigma(1), \dots, \sigma(m-1)\}, X_{u_1+\dots+u_m} = \sigma(m).
\end{aligned}$$

Ainsi, l'événement $\{U_1 = u_1, \dots, U_m = u_m\}$ est la réunion disjointe

$$\biguplus_{\sigma \in S_m} \Big\{ \begin{aligned}
&X_{u_1} = \sigma(1), X_{u_1+1} = \sigma(1), \dots,\\
&X_{u_1+u_2-1} = \sigma(1), X_{u_1+u_2} = \sigma(2), X_{u_1+u_2+1} \in \{\sigma(1), \sigma(2)\}, \dots,\\
&X_{u_1+u_2+u_3-1} \in \{\sigma(1), \sigma(2)\}, X_{u_1+u_2+u_3} = \sigma(3), \dots\\
&\dots, X_{u_1+\dots+u_m-1} \in \{\sigma(1), \dots, \sigma(m-1)\}, X_{u_1+\dots+u_m} = \sigma(m) \Big\}.
\end{aligned}$$

Les événements cette réunion sont de même probabilité, il y en a $m\,!$, d'où la probabilité de la réunion lorsque $u_1 = 1$:

$$\begin{aligned}
m\,! \times \Big(\frac{1}{m}\Big)^{u_2} \times \frac{1}{m}\Big(\frac{2}{m}\Big)^{u_3-1} \times \frac{1}{m}\Big(\frac{3}{m}\Big)^{u_4-1} \times \dots \times \frac{1}{m}\Big(\frac{m-1}{m}\Big)^{u_m-1} \times \frac{1}{m}\\
= \frac{m-1}{m}\Big(\frac{1}{m}\Big)^{u_2-1} \times \frac{m-2}{m}\Big(\frac{2}{m}\Big)^{u_3-1} \times \frac{m-3}{m}\Big(\frac{3}{m}\Big)^{u_4-1} \times \dots \times \frac{1}{m}\Big(\frac{m-1}{m}\Big)^{u_m-1}
\end{aligned}$$

(lorsque $u_1 = 0$, cette probabilité est nulle). Pour $1 \leqslant k \leqslant m$, en sommant pour tout $j \neq k$ sur u_j de 1 à l'infini, on constate que

$$\mathbb{P}\{U_k = u_k\} = \frac{m-k+1}{m}\Big(\frac{k-1}{m}\Big)^{u_k-1}$$

pour tout $u_k \in \mathbb{N}^*$, et donc que la loi de U_k est la loi géométrique de paramètre $(m-k+1)/m$. (Ceci est encore valable pour $k = 1$.)

(iii) La loi conjointe de $U_1, \dots, U_m$ est, d'après ce que nous avons obtenu, la loi produit des lois des U_k. Ces variables sont donc indépendantes.

(iv) Pour $1 \leqslant k \leqslant m$, on a

$$G_k(s) = \mathbb{E}\big[s^{U_k}\big] = \sum_{n \geqslant 1} s^n\, \mathbb{P}\{U_k = n\} = \frac{p_k s}{1-(1-p_k)s} = \frac{(m-k+1)s}{m-(k-1)s}$$

avec $p_k = (m-k+1)/m$, et ainsi, par indépendance

$$G_T(s) = G_{U_1+\dots+U_m}(s) = \prod_{k=1}^{m} G_k(s) = \dots = s^m \prod_{k=2}^{m} \frac{m-k+1}{m-(k-1)s}.$$

Et, simplement par linéarité,

$$\begin{aligned}
\mathbb{E}[T] &= \mathbb{E}[U_1 + \dots + U_m] = \mathbb{E}[U_1] + \dots + \mathbb{E}[U_m]\\
&= 1 + \frac{m-1}{m} + \frac{m-2}{m} + \dots + \frac{1}{m} = m\Big(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{m}\Big).
\end{aligned}$$

Exercice 16 (contrôle des naissances). — Soit $X = (X_n)_{n \geqslant 1}$ une suite de variables aléatoires indépendantes suivant chacune la loi de Bernoulli symétrique $\mathcal{B}(1, 1/2)$. Les valeurs obtenues correspondent à la naissance au n-ième essai d'un garçon ou d'une fille, et on suppose

qu'il n'y a pas de naissance multiple. Les parents se fixent certains objectifs ce qui conduit à temps d'arrêt N qui est le nombre d'enfants obtenus.

D'après l'identité de Wald, quels que soient les objectifs choisis par les parents, la moyenne de garçons (ou de filles) obtenus est $\mathbb{E}[N]/2$, et donc, ne dépend que de l'espérance du temps d'arrêt N.

(i) La famille décide d'arrêter de procréer dès qu'elle a obtenu deux garçons consécutivement. Évaluer le nombre moyen d'enfants et de garçons dans la famille.

(ii) La famille décide d'arrêter de procréer dès qu'elle a obtenu un garçon succédant immédiatement à une fille. Évaluer le nombre moyen d'enfants et de garçons dans la famille.

Correction. — (i) Soit S le premier instant où on obtient deux garçons consécutifs. En décomposant l'espérance de S sur les événements $\{X_1 = 1\}$ et $\{X_1 = 0\}$, on a

$$\begin{aligned}\mathbb{E}[S] &= \mathbb{P}\{X_1 = 1\} \times \mathbb{E}[S \,|\, X_1 = 1] + \mathbb{P}\{X_1 = 0\} \times \mathbb{E}[S \,|\, X_1 = 0] \\ &= \big(\mathbb{E}[S \,|\, X_1 = 1] + \mathbb{E}[S + 1]\big)/2 = \big(\mathbb{E}[S \,|\, X_1 = 1] + \mathbb{E}[S] + 1\big)/2\end{aligned}$$

puisque si on a une fille pour commencer, on a fait un essai pour rien. On a de même

$$\begin{aligned}\mathbb{E}[S \,|\, X_1 = 1] &= \mathbb{P}\{X_2 = 1\} \times \mathbb{E}[S \,|\, X_1 = 1, X_2 = 1] + \mathbb{P}\{X_2 = 0\} \times \mathbb{E}[S \,|\, X_1 = 1, X_2 = 0] \\ &= \big(2 + \mathbb{E}[S + 2]\big)/2 = 2 + \mathbb{E}[S]/2,\end{aligned}$$

et ce, pour des raisons identiques. On obtient ainsi

$$\mathbb{E}[S] = \big(2 + \mathbb{E}[S]/2 + 1 + \mathbb{E}[S]\big)/2 = 3/2 + 3/4\,\mathbb{E}[S],$$

d'où $\mathbb{E}[S] = 6$.

(ii) Soit T le premier instant où on obtient une fille précédée d'une garçon. En décomposant l'espérance de T sur les événements $\{X_1 = 1\}$ et $\{X_1 = 0\}$, on a

$$\begin{aligned}\mathbb{E}[T] &= \mathbb{P}\{X_1 = 1\} \times \mathbb{E}[T \,|\, X_1 = 1] + \mathbb{P}\{X_1 = 0\} \times \mathbb{E}[T \,|\, X_1 = 0] \\ &= \big(\mathbb{E}[T \,|\, X_1 = 1] + \mathbb{E}[T + 1]\big)/2 = \big(\mathbb{E}[T \,|\, X_1 = 1] + \mathbb{E}[T] + 1\big)/2\end{aligned}$$

puisque si on a une fille pour commencer, on a fait un essai pour rien. On a de même

$$\begin{aligned}\mathbb{E}[T \,|\, X_1 = 1] &= \mathbb{P}\{X_2 = 1\} \times \mathbb{E}[T \,|\, X_1 = 1, X_2 = 1] + \mathbb{P}\{X_2 = 0\} \times \mathbb{E}[T \,|\, X_1 = 1, X_2 = 0] \\ &= \big(1 + \mathbb{E}[T \,|\, X_1 = 1] + 2\big)/2 = 3/2 + \mathbb{E}[T \,|\, X_1 = 1]/2,\end{aligned}$$

donc $\mathbb{E}[T \,|\, X_1 = 1] = 3$ et $\mathbb{E}[T] = (3 + 1 + \mathbb{E}[T])/2$, soit $\mathbb{E}[T] = 4$.

Remarque. — Ceci anticipe sur les chaînes de Markov, et devrait être traité plus rigoureusement.

FEUILLE D'EXERCICES N° 2. — ESPÉRANCES CONDITIONNELLES ET MARTINGALES À VALEURS DANS $\mathbb{R}^D$

EXERCICE 1. — Parmi 100 questions possibles pour un questionnaire à choix multiples de 20 questions, un étudiant en révise une proportion p (si $p < 1$ ce n'est pas un bon étudiant). Aux questions qu'il a révisé, il répond correctement à coup sûr. Pour les autres, il répond *au hasard* parmi les 4 choix possibles. Calculer l'espérance de sa note. (*Indication.* On posera X le nombre de questions parmi les 20 choisies connues par l'étudiant, et $N = Y + X$ sa note totale. On déterminera la loi de Y sachant X ce qui permettra de calculer $\mathbb{E}[N \mid X]$ et ainsi $\mathbb{E}[X]$. On rappelle que l'espérance d'une variable aléatoire de loi la loi $\mathcal{H}(N, n, p)$ est np.)

Correction. — Soit X le nombre de questions connues par l'étudiant dans les 20 questions tirées parmi les 100 questions au total. La loi de X est alors la loi hypergéométrique $\mathcal{H}(100, 20, p)$. Soit N sa note et $Y = N - X$ qui est le nombre de questions parmi celles que n'avait pas révisées l'étudiant mais auxquelles il a répondu juste. Pour $k \in \{0, \dots, 20\}$, on a $\mathbb{E}[Y \mid X = k] = (20 - k)/4 = 5 - k/4$ — l'étudiant ne répond correctement en moyenne qu'à un quart des questions qu'il ne connaît pas —, donc $\mathbb{E}[Y \mid X] = 5 - X/4$. Alors

$$\begin{aligned}\mathbb{E}[N] &= \mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] \\ &= \mathbb{E}[X] + \mathbb{E}\big[\mathbb{E}[Y \mid X]\big] = \mathbb{E}[X] + \mathbb{E}[5 - X/4] \\ &= 5 + 3\,\mathbb{E}[X]/4 = 5 + 3 \times 20p/4 = 5 + 15p.\end{aligned}$$

Remarque. — Ce résultat est cohérent : si $p = 1$, $\mathbb{E}[N] = 20$, alors que si $p = 0$, $\mathbb{E}[N] = 5$. C'est un exercice qui fait réfléchir…

EXERCICE 2. — Un mineur est au fond d'une mine comportant trois portes. La première porte donne sur un tunnel qui le conduira à l'extérieur après trois heures de marche. La deuxième porte donne sur un tunnel qui le ramènera au fond de la mine après cinq heures de marche. La troisième porte donne sur un tunnel qui le ramènera au fond de la mine après sept heures de marche.

Si on suppose que le mineur choisit l'une des portes avec probabilité 1/3 à chaque passage, qu'elle est l'espérance du temps de parcours avant d'atteindre l'extérieur ?

Correction. — Soit T le temps que mettra le mineur à sortir du fond de la mine. Soit C la variable qui indique la porte choisie au premier passage. On a

$$\mathbb{E}[T] = \mathbb{P}\{C = 1\}\,\mathbb{E}[T \mid C = 1] + \mathbb{P}\{C = 2\}\,\mathbb{E}[T \mid C = 2] + \mathbb{P}\{C = 3\}\,\mathbb{E}[T \mid C = 3].$$

Évidemment, $\mathbb{E}[T \mid C = 1] = 3$, et nous affirmons que $\mathbb{E}[T \mid C = 2] = 5 + \mathbb{E}[T]$ et $\mathbb{E}[T \mid C = 3] = 7 + \mathbb{E}[T]$ car revenant au point de départ le mineur choisira à nouveau une porte de la même manière qu'au premier essai. Ainsi,

$$\mathbb{E}[T] = \frac{1}{3}3 + \frac{1}{3}(5 + \mathbb{E}[T]) + \frac{1}{3}(7 + \mathbb{E}[T]), \qquad \text{c'est-à-dire} \qquad \mathbb{E}[T] = 15.$$

EXERCICE 3. — Soit $(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{F})$ un espace probabilisé filtré.

(i) Montrer que pour tout $A \in \mathcal{A}$, $X_n = \mathbb{E}[\mathbb{1}_A \mid \mathcal{F}_n] = \mathbb{P}(A \mid \mathcal{F}_n)$ définit une martingale. Est-elle convergente ? Si oui, que dire de sa limite (noter qu'*a priori* on a seulement $\mathcal{F}_\infty \subset \mathcal{A}$) ?

(ii) Montrer que pour tout $Y \in L^1(\Omega, \mathcal{A}, \mathbb{P})$, $X_n = \mathbb{E}[Y \,|\, \mathcal{F}_n]$ définit une martingale. Est-elle convergente ? Si oui, que dire de sa limite ?

Correction. — (i) Pour chaque n, nous choisissons un représentant X_n de $\mathbb{P}(A \,|\, \mathcal{F}_n)$ qui soit $\mathcal{F}_n$-mesurable (pas d'augmentation particulière). Soit n donné. Par positivité de l'espérance conditionnelle, X_n est une variable aléatoire positive et on a $\mathbb{E}[X_n] = \mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$ qui est fini. Donc X_n est intégrable. Pour $m \leqslant n$, nous avons par transitivité de l'espérance conditionnelle $X_m = \mathbb{E}[\mathbb{1}_A \,|\, \mathcal{F}_m] = \mathbb{E}[\mathbb{E}[\mathbb{1}_A \,|\, \mathcal{F}_n] \,|\, \mathcal{F}_m] = \mathbb{E}[X_n \,|\, \mathcal{F}_m]$ presque sûrement puisque $\mathcal{F}_m \subset \mathcal{F}_n$. Ainsi X est une martingale dans la filtration $\mathcal{F}$.

La martingale X est bornée dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$. Nous savons donc qu'il existe X_∞ une variable aléatoire $\mathcal{F}_\infty$-mesurable telle que X converge presque sûrement vers X_∞ (convergence des martingales, cas L^1, (i)).

Les propositions *a*), *b*) et *c*) de (ii) du théorème de convergence des martingales, cas L^1, sont satisfaites puisque *d*) est vraie par définition de la martingale X. Il faut cependant prendre garde que les « X_∞ » du (i) et du (ii) ne sont pas nécessairement les mêmes. Ainsi, X est uniformément intégrable, convergente dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$ et même $L^1(\Omega, \mathcal{F}_\infty, \mathbb{P})$. Si X_∞ est la limite presque sûre de (i), alors on a $X_\infty \in L^1(\Omega, \mathcal{F}_\infty, \mathbb{P})$ et $X_n = \mathbb{E}[X_\infty \,|\, \mathcal{F}_n]$. Il se peut que $X_\infty \neq \mathbb{1}_A$, en particulier si $\mathcal{F}_\infty \subsetneq \mathcal{A}$.

(ii) Pour chaque n, nous choisissons un représentant X_n de $\mathbb{E}[Y \,|\, \mathcal{F}_n]$ qui soit $\mathcal{F}_n$-mesurable (pas d'augmentation particulière). Soit n donné. L'espérance conditionnelle étant une contraction dans L^1, X_n est une variable aléatoire intégrable comme on peut le voir en intégrant l'inégalité de Jensen conditionnelle

$$\mathbb{E}[|X_n|] = \mathbb{E}\big[\big|\mathbb{E}[Y \,|\, \mathcal{F}_n]\big|\big] \leqslant \mathbb{E}\big[\mathbb{E}[|Y| \,|\, \mathcal{F}_n]\big] = \mathbb{E}[|Y|]$$

(la suite $(X_n)_n$ est même bornée dans L^1). Pour $m \leqslant n$, nous avons par transitivité de l'espérance conditionnelle

$$X_m = \mathbb{E}[Y \,|\, \mathcal{F}_m] = \mathbb{E}[\mathbb{E}[Y \,|\, \mathcal{F}_n] \,|\, \mathcal{F}_m] = \mathbb{E}[X_n \,|\, \mathcal{F}_m] \qquad \text{presque sûrement}$$

puisque $\mathcal{F}_m \subset \mathcal{F}_n$. Ainsi $(X_n)_n$ est une martingale dans la filtration $\mathcal{F}$.

La martingale $(X_n)_n$ est bornée dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$ comme on l'a vu plus haut. Nous savons donc qu'il existe X_∞ une variable aléatoire $\mathcal{F}_\infty$-mesurable telle que $(X_n)_n$ converge presque sûrement vers X_∞ (convergence des martingales, cas L^1, (i)).

Les propositions *a*), *b*) et *c*) de (ii) du théorème de convergence des martingales, cas L^1, sont satisfaites puisque *d*) est vraie par définition de la martingale $(X_n)_n$. Il faut cependant prendre garde que les « X_∞ » du (i) et du (ii) ne sont pas nécessairement les mêmes. Ainsi, $(X_n)_n$ est uniformément intégrable, convergente dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$ et même $L^1(\Omega, \mathcal{F}_\infty, \mathbb{P})$. Si X_∞ est la limite presque sûre de (i), alors on a $X_\infty \in L^1(\Omega, \mathcal{F}_\infty, \mathbb{P})$ et $X_n = \mathbb{E}[X_\infty \,|\, \mathcal{F}_n]$. Il se peut que $X_\infty \neq Y$, en particulier si $\mathcal{F}_\infty \subsetneq \mathcal{A}$.

Exercice 4. — Soient $(\xi_n)_{n\geqslant 1}$ une suite de variables aléatoires réelles indépendantes, identiquement distribuées et intégrables. On pose $X_n = \sum_{k=1}^n \xi_k$ pour $n \geqslant 1$ et $X_0 = 0$. La filtration $\mathcal{F}$ considérée étant la filtration naturelle de $(\xi_n)_{n\geqslant 1}$ (avec $\mathcal{F}_0 = \{\varnothing, \Omega\}$), que dire du processus X en fonction de $m = \mathbb{E}[\xi_1]$?

Correction. — Ce que nous considérons est appelé « marche aléatoire » (homogène). Remarquons que les filtrations naturelles $\mathcal{F}^\xi$ et $\mathcal{F}^X$ sont identiques. Pour tout n, X_n est une somme finie de variables aléatoires intégrables et est donc intégrable. Pour $n \geqslant 0$, on a $\mathbb{E}[X_{n+1} \,|\, \mathcal{F}_n] = \mathbb{E}[X_n + \xi_{n+1} \,|\, \mathcal{F}_n] = \mathbb{E}[X_n \,|\, \mathcal{F}_n] + \mathbb{E}[\xi_{n+1} \,|\, \mathcal{F}_n] = X_n + m$ presque sûrement puisque ξ_{n+1} est indépendante de $\mathcal{F}_n = \sigma\{\xi_1, \ldots, \xi_n\}$. Il est alors clair que

- $m \leqslant 0 \iff X$ est une sur-martingale ;
- $m = 0 \iff X$ est une martingale ;

- $m \geqslant 0 \iff X$ est une sous-martingale.

Pendant que nous y sommes, nous pouvons préciser un peu le comportement asymptotique de ce processus. Par la loi forte des grands nombres de Kolmogorov, nous savons que X_n/n tend presque sûrement vers m lorsque n tend vers l'infini. Ainsi, si $m \neq 0$, nous avons presque sûrement, $X_n \sim n \times m$ quand n tend vers l'infini.

Pour $m = 0$, lorsque $\sigma^2 = \mathrm{Var}(\xi)$ est fini et non nul, le théorème central limite affirme que les lois des $\frac{X_n}{\sqrt{n}\sigma}$ converge vers la loi $\mathcal{N}(0,1)$. Notamment, X_n ne converge en aucun sens habituel (presque sûrement, en probabilité, dans L^p, puisque si c'était le cas, $\frac{X_n}{\sqrt{n}\sigma}$ convergerait vers 0 au même sens, et alors la suite des lois convergerait vers $\delta_{\{0\}}$.

Toujours pour $m = 0$, si $\sigma^2 = 0$, alors $\xi_n = \mathrm{Cte} = 0$ presque sûrement et ainsi $X_n = 0$ presque sûrement, qui converge vers 0.

Enfin, si $m = 0$ et $\sigma = \infty$, la « non convergence » est encore plus forte que dans le cas $\sigma > 0$. Il faudrait évoquer des résultats plus fins pour préciser tout ça.

Exercice 5. — Soit $(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{F})$ un espace probabilisé filtré où le temps est $\mathbb{T} = \mathbb{N}$. En posant (avec $\mathcal{F}_{-1} = \{\varnothing, \Omega\}$), on dit qu'un processus $H = (H_n)_{n\geqslant 0}$ est prévisible si pour tout $n \geqslant 0$, H_n est $\mathcal{F}_{n-1}$-mesurable. Montrer que si X est une martingale et H un processus prévisible localement borné (ce qui signifie ici que chaque H_n est une variable aléatoire bornée), alors

$$Y_0 = 0, \qquad Y_n = \sum_{k=1}^{n} H_k \times (X_k - X_{k-1}) = \sum_{k=1}^{n} H_k \Delta X_k, \quad n \geqslant 1,$$

définit une martingale.

Correction. — Soient $\Delta X_k = X_k - X_{k-1}$, $k \geqslant 1$. Ces accroissements sont ceux d'une martingale : ils sont intégrables et satisfont $\mathbb{E}[\Delta X_k \,|\, \mathcal{F}_\ell] = 0$ presque sûrement pour tous $k > \ell$.

Soit H un processus prévisible localement borné (H_k est $\mathcal{F}_{k-1}$-mesurable bornée). Alors $H_k \Delta X_k$ est un accroissement de martingale car c'est intégrable et

$$\mathbb{E}[H_k \Delta X_k \,|\, \mathcal{F}_\ell] = \mathbb{E}\big[\mathbb{E}[H_k \Delta X_k \,|\, \mathcal{F}_{k-1}] \,\big|\, \mathcal{F}_\ell\big] = \mathbb{E}\big[H_k\, \mathbb{E}[\Delta X_k \,|\, \mathcal{F}_{k-1}] \,\big|\, \mathcal{F}_\ell\big] = \mathbb{E}[H_k \times 0 \,|\, \mathcal{F}_\ell] = 0$$

presque sûrement pour $k > \ell$. Puisque $Y_0 = 0$, Y_n est intégrable puisque somme finie de variables aléatoires intégrables, et c'est une martingale d'après ce qui vient d'être vu de ses accroissements.

Remarque. — Cette manière d'obtenir une martingale à partir d'une autre est souvent appelée *transformée de Burkhölder*, c'est en fait la version discrète de l'intégrale stochastique par rapport à une martingale.

Exercice 6 (martingale de Wald). — Soient $(\xi_n)_{n\geqslant 1}$ une suite de variables aléatoires réelles indépendantes, identiquement distribuées admettant une fonction génératrice $G(u) = \mathbb{E}[\mathrm{e}^{u\xi_n}] \in]0, \infty[$ pour au moins un réel $u \neq 0$.

(i) Montrer que

$$X_0 = 1, \qquad X_n = G(u)^{-n}\, \mathrm{e}^{u(\xi_1 + \cdots + \xi_n)}, \qquad n \geqslant 1,$$

définit une martingale.

(ii) Préciser pourquoi cette martingale est presque sûrement convergente.

(iii) Que se passe-t-il si $u = 0$ ou ξ est presque sûrement constante ?

(iv) Qu'en est-il si aucune des deux conditions précédentes est satisfaite ?

Correction. — (i) Nous avons $X_{n+1} = X_n \times \mathrm{e}^{u\xi_{n+1}}/G(u)$. Toutes les variables aléatoires sont positives et $\mathrm{e}^{u\xi_{n+1}}/G(u)$ est indépendante de $\mathcal{F}_n$. Ainsi,

$$\mathbb{E}[X_{n+1} \,|\, \mathcal{F}_n] = \mathbb{E}\left[X_n \times \frac{\mathrm{e}^{u\xi_{n+1}}}{G(u)} \,\middle|\, \mathcal{F}_n\right] = X_n \times \frac{\mathbb{E}\big[\mathrm{e}^{u\xi_{n+1}} \,\big|\, \mathcal{F}_n\big]}{G(u)} = X_n \times \frac{\mathbb{E}\big[\mathrm{e}^{u\xi_{n+1}}\big]}{G(u)} = X_n$$

presque sûrement. Il est de plus clair que chaque X_n est intégrable et $\mathbb{E}[X_n] = 1$.

(ii) On peut se demander quel est le comportement asymptotique de cette martingale positive. Un théorème fondamental est le théorème de convergence des sur-martingales positives : toute sur-martingale positive converge presque sûrement. Ceci peut s'interpréter comme une généralisation probabiliste de la convergence des suites numériques, décroissantes, minorées. De toutes façons, X est aussi une martingale bornée dans L^1. Notons X_∞ sa limite presque sûre.

Constatons que $\ln X_n = (u\xi_1 - \ln G(u)) + \cdots + (u\xi_n - \ln G(u))$ qui est une marche aléatoire à accroissements intégrables. Son comportement asymptotique est en premier lieu déterminé par $m = \mathbb{E}[u\xi - \ln G(u)] = \mathbb{E}[u\xi] - \ln G(u) = \mathbb{E}[u\xi] - \ln(\mathbb{E}[\exp(u\xi)])$. La fonction logarithme étant concave et les espérances étant bien définies, l'inégalité de Jensen implique que $\mathbb{E}[u\xi] \leqslant \ln(\mathbb{E}[\exp(u\xi)])$, et donc $m \leqslant 0$. De plus, par stricte concavité du logarithme, on ne peut avoir égalité que si $\exp(u\xi)$ est presque sûrement constante, c'est-à-dire $u = 0$ ou ξ variable aléatoire constante.

(iii) Si $u = 0$ ou ξ constante, il est alors clair que $X_n = 1$ presque sûrement pour tout n, et donc $X_\infty = 1$.

(iv) Mais pour $u \neq 0$ et ξ non constante, on a $m < 0$, et $\ln X_n \sim n \times m$, soit $X_n \sim \exp(n \times m) \to X_\infty = 0$ quand n tend vers l'infini. Nous avons alors un exemple de martingale convergente presque sûrement qui ne converge pas dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$.

Exercice 7 (rapport de vraisemblance). — Soit μ une mesure de probabilité sur $\mathbb{R}$ absolument continue par rapport à la mesure de Lebesgue dont la densité est ou bien $f : \mathbb{R} \to \mathbb{R}_+$ ou bien $g : \mathbb{R} \to \mathbb{R}_+$. Le problème est de déterminer laquelle de ces fonctions est la (une) densité de μ. Pour cela on se donne une suite $X = (X_n)_{n\geqslant 1}$ de variables aléatoires de loi μ et on considère

$$M_n = \prod_{k=1}^{n} \frac{g(X_k)}{f(X_k)}, \qquad n \geqslant 1,$$

où on aura convenu que $0/0 = 0$. Le processus $M = (M_n)_{n\geqslant 0}$ ($M_0 = 1$) est appelé « rapport de vraisemblance » ; il est à valeurs dans $\mathbb{R}_+ \cup \{\infty\}$.

(i) Montrer que si la densité de μ est f alors le processus M est presque sûrement à valeurs finies.

(ii) Montrer que si la densité de μ est f, alors le processus M est une martingale. En déduire que M est presque sûrement convergent.

(iii) Si de plus $f \neq g$ (sur un ensemble de mesure de Lebesgue non nulle), quelle est la variable aléatoire limite de M ?

(iv) De ce qui précède, que devrait-on observer si la densité de μ est en réalité g ?

Correction. — Non disponible.

Exercice 8 (variation quadratique). — Soit $X = (X_n)_{n\geqslant 0}$ une martingale réelle telle que $\mathbb{E}[X^2] < \infty$ pour tout n.

(i) Que dire de $X^2 = (X_n^2)_{n\geqslant 0}$?

(ii) On définit la variation quadratique (crochet droit) de X par

$$[X, X]_n = \sum_{k=1}^{n} (X_k - X_{k-1})^2 = \sum_{k=1}^{n} \Delta X_k^2, \quad \text{pour } n \geqslant 1, \qquad \text{et} \qquad [X, X]_0 = 0.$$

Que peut-on dire du processus $X^2 - [X, X]$?

(iii) On définit la variation quadratique prévisible (crochet oblique) de X par

$$\langle X, X\rangle_n = \sum_{k=1}^{n} \mathbb{E}[(X_k - X_{k-1})^2 \,|\, \mathcal{F}_{k-1}] = \sum_{k=1}^{n} \mathbb{E}[\Delta X_k^2 \,|\, \mathcal{F}_{k-1}], \quad \text{pour } n \geqslant 1,$$

et $\langle X, X\rangle_0 = 0$. Que peut-on dire du processus $X^2 - \langle X, X\rangle$?

(iv) Exprimer les variations quadratiques d'une transformée de Burkhölder.

Correction. — Non disponible.

Université de Poitiers
Département de Mathématiques

FEUILLE D'EXERCICES N° 3. — PROCESSUS DE MARKOV, GÉNÉRALITÉS

Exercice 1. — (i) Montrer que le produit de deux matrices stochastiques est une matrice stochastique.

(ii) Montrer qu'une combinaison barycentrique de matrices stochastiques est une matrice stochastique.

(iii) Montrer que la matrice d'une permutation est bistochastique.

Correction. — Rappelons qu'une matrice stochastique est une matrice carrée

$$P = \big(P(x,y)\big)_{x,y\in E}$$

à coefficients positifs telle que pour tout $x \in E$, $\sum_{y\in E} P(x,y) = 1$ (les sommes des coefficients sur chaques lignes sont égales à 1).

(i) Soient P, Q deux matrices stochastiques sur E. Le coefficient en (x,z) de leur produit PQ est

$$PQ(x,z) = \sum_{y\in E} P(x,y)Q(y,z)$$

qui est positif puisque tous les coefficients $P(x,y)$ et $Q(y,z)$ sont positifs. On a

$$\begin{aligned}\sum_{z\in E} PQ(x,z) = \sum_{z\in E}\sum_{y\in E} P(x,y)Q(y,z) &= \sum_{y\in E}\sum_{z\in E} P(x,y)Q(y,z)\\ &= \sum_{y\in E} P(x,y)\sum_{z\in E} Q(y,z) = \sum_{y\in E} P(x,y)\times 1 = \sum_{y\in E} P(x,y) = 1,\end{aligned}$$

les deux dernières égalités résultant que les sommes par lignes de P et de Q sont égales à 1. La deuxième égalité consiste simplement en une interversion des ordres de sommation.

(ii) Une combinaison barycentrique de deux matrices stochastiques P et Q s'écrit $tP+(1-t)Q$ avec $t \in [0,1]$. Pour $x, y \in E$, $(tP + (1-t)Q)(x,y) = tP(x,y) + (1-t)Q(x,y) \geqslant 0$ puisque somme de deux nombres positifs. On a enfin, pour $x \in E$, $\sum_{y\in E}(tP + (1-t)Q)(x,y) = t\sum_{y\in E} P(x,y) + (1-t)\sum_{y\in E} Q(x,y) = t + (1-t) = 1$. Les combinaisons barycentriques préservent les matrices stochastiques.

Remarque. — Plus généralement, il est facile de voir qu'une combinaison barycentrique de noyaux de transition est encore un noyau de transition.

(iii) Soit $\sigma \in \mathfrak{S}_n$ une permutation de $\{1,\dots,n\}$. Nous pensons à σ comme décrivant les transitions possibles, ainsi on transite de i à $j \in \{1,\dots,n\}$ si et seulement si $j = \sigma(i)$. La matrice de transition P_σ associée est alors donnée par les coefficients $P_\sigma(x,y)$ probabilités de transiter de i à j, ainsi $P_\sigma(x,y) = \delta(\sigma(i), j)$ (symbole de Kronecker valant 1 si les deux arguments sont égaux, 0 sinon). La ligne $i \in \{1,\dots,n\}$ de P_σ ne comporte qu'un 1 et des 0, ses coefficients sont donc positifs de somme égale à 1.

Une matrice est bistochastique si elle est stochastique et si les sommes de ses coefficients en colonnes sont égales à 1, autrement dit, si sa transposée est aussi stochastique. On a

$${}^tP_\sigma = t\big(\delta(\sigma(i),j)\big)_{i,j=1}^n = \big(\delta(\sigma(j),i)\big)_{i,j=1}^n = \big(\delta(i,\sigma(j))\big)_{i,j=1}^n = \big(\delta(\sigma^{-1}(i),j)\big)_{i,j=1}^n = P_{\sigma^{-1}}$$

puisque σ est une bijection. Ainsi ${}^tP_\sigma = P_{\sigma^{-1}}$ qui est la matrice de la permutation σ^{-1} et qui est donc elle aussi stochastique. Finalement, la matrice d'une permutation est bistochastique.

Évidemment, la formule définissant P_σ indique que sur la colonne j, il existe un unique numéro de ligne i (tel que $\sigma(i) = j$, ou encore $i = \sigma^{-1}(j)$) pour lequel le coefficient correspondant est égal à 1, les autres étant nuls. Ceci assure que les sommes par colonnes sont égales à 1.

Exercice 2. — Considérons 2 urnes contenant au total N boules. Une boule est choisie au hasard et transvasée dans l'autre urne. Décrire ce mécanisme à l'aide d'une matrice stochastique en s'intéressant, par exemple, au nombre de boules contenues dans l'une des deux urnes.

Correction. — Si l'urne A contient i boules, l'urne B en contient $N-i$. Tirer une boule dans l'urne A a alors probabilité i/N et fait passer à l'état $i-1$. Tirer une boule dans l'urne B a alors probabilité $(N-i)/N$ et fait passer à l'état $i+1$. Ainsi, pour $0 \leqslant i, j \leqslant N$,

$$P(i,j) = \begin{cases} i/N & \text{si } j = i-1, \\ (N-i)/N & \text{si } j = i+1, \\ 0 & \text{sinon,} \end{cases}$$

soit :

$$P = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ \frac{1}{N} & 0 & \frac{N-1}{N} & \ddots & & & \vdots \\ 0 & \frac{2}{N} & 0 & \frac{N-2}{N} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \frac{N-2}{N} & 0 & \frac{2}{N} & 0 \\ \vdots & & & \ddots & \frac{N-1}{N} & 0 & \frac{1}{N} \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Exercice 3. — Le temps est supposé discret. Un processus X est markovien dans la filtration $\mathcal{F}$ si et seulement si pour tous $0 \leqslant n < m$, et toute application $f : (E, \mathcal{E}) \to \mathbb{R}$ bornée, on a

$$\mathbb{E}[f(X_m) \,|\, \mathcal{F}_n] = \mathbb{E}[f(X_m) \,|\, X_n] \qquad \text{presque sûrement,} \tag{A}$$

où le second conditionnement est effectué par rapport à la tribu engendrée par X_n.

Montrer que la condition précédente équivaut à : pour tous $n \geqslant 0$, et toute application $f : E \to \mathbb{R}$ bornée, on a

$$\mathbb{E}[f(X_{n+1}) \,|\, \mathcal{F}_n] = \mathbb{E}[f(X_{n+1}) \,|\, X_n] \qquad \text{presque sûrement.} \tag{B}$$

(*Indication.* — Pour le sens [B] implique [A], on pourra utiliser le fait que si Z est une variable aléatoire intégrable [bornée par exemple], alors une version de $\mathbb{E}[Z \,|\, X_k]$ peut s'écrire $f_k(X_k)$ [lemme de Doob].)

Correction. — On a [A] implique [B] en prenant $m = n+1$ pour $n \geqslant 0$. Réciproquement supposons que [B] ait lieu. Soit $f : E \to \mathbb{R}$ une application bornée, et $0 \leqslant n < m$. Puisque $\mathcal{F}_{m-1} \supset \mathcal{F}_n$ comme $m - 1 \geqslant n$, on a (presque sûrement)

$$\begin{aligned} \mathbb{E}[f(X_m) \,|\, \mathcal{F}_n] &= \mathbb{E}\big[\mathbb{E}[f(X_m) \,|\, \mathcal{F}_{m-1}] \,\big|\, \mathcal{F}_n\big] \\ &= \mathbb{E}\big[\mathbb{E}[f(X_m) \,|\, X_{m-1}] \,\big|\, \mathcal{F}_n\big] \qquad \text{par la propriété [B]} \\ &= \mathbb{E}[f_{m-1}(X_{m-1}) \,|\, \mathcal{F}_n\big] \qquad \text{par le lemme de Doob.} \end{aligned}$$

En itérant, on obtient $\mathbb{E}[f(X_m) \,|\, \mathcal{F}_n] = \mathbb{E}[f_{n+1}(X_{n+1}) \,|\, \mathcal{F}_n] = f_n(X_n)$ presque sûrement, qui est $\mathcal{F}_n$-mesurable. Par conséquent, cette espérance conditionnelle est égale à $\mathbb{E}[f(X_m) \,|\, X_n]$

puisqu'alors

$$\begin{aligned}\mathbb{E}[f(X_m)\,|\,\mathcal{F}_n] &= \mathbb{E}\big[\mathbb{E}[f(X_m)\,|\,\mathcal{F}_n]\;\big|\;X_n\big] \quad \text{puisque le terme de gauche est } \sigma(X_n)\text{-mesurable}\\ &= \mathbb{E}[f(X_m)\,|\,X_n] \qquad \text{puisque } \mathcal{F}_n \supset \sigma(X_n).\end{aligned}$$

On a donc bien $[B]$ implique $[A]$.

Exercice 4. — Soient $X_0, \xi_1, \ldots, \xi_n, \ldots$ des variables aléatoires indépendantes à valeurs dans $\mathbb{R}^d$. Pour $n \geqslant 1$, on pose $X_n = X_0 + \xi_1 + \cdots + \xi_n$, et on considère ce processus dans sa filtration naturelle.

(i) Montrer que le processus X est markovien.

(ii) Lorsque les $(\xi_n)_{n\geqslant 1}$ sont identiquement distribuées, montrer que le processus X est homogène et déterminer son semi-groupe de transition.

Remarque. — On pourra se contenter de considérer des variables aléatoires à valeurs dans $\mathbb{Z}^d$ avec $d = 1$.

Correction. — Nous allons supposer que les variables $X_0, \xi_1, \ldots, \xi_n, \ldots$ sont à valeurs dans un groupe E (plus généralement, un ensemble stable par addition, voire un ensemble muni d'une opération interne $\top : E \times E \to E$ ni nécessairement commutative, ni nécessaire associative) au plus dénombrable de $\mathbb{R}$. Ceci va considérablement simplifier les conditionnements tout en donnant l'intuition de ce qu'il se passe.

(i) Soit $f : E \to \mathbb{R}$ une fonction mesurable bornée. Pour $n \geqslant 0$ donné, écrivons $1 = \sum_{x\in E} \mathbb{1}_{\{X_n=x\}}$. Nous avons alors

$$\begin{aligned}\mathbb{E}[f(X_{n+1})\,|\,\mathcal{F}_n] &= \mathbb{E}[f(X_n+\xi_{n+1})\,|\,\mathcal{F}_n]\\ &= \mathbb{E}\left[\sum_{x\in E} \mathbb{1}_{\{X_n=x\}} f(X_n+\xi_{n+1}) \;\middle|\; \mathcal{F}_n\right]\\ &= \mathbb{E}\left[\sum_{x\in E} \mathbb{1}_{\{X_n=x\}} f(x+\xi_{n+1}) \;\middle|\; \mathcal{F}_n\right]\\ &= \sum_{x\in E} \mathbb{1}_{\{X_n=x\}}\, \mathbb{E}[f(x+\xi_{n+1})\,|\,\mathcal{F}_n]\end{aligned}$$

puisque $\mathbb{1}_{\{X_n=x\}}$ est $\mathcal{F}_n$-mesurable. De plus, pour tout $x \in E$, $f(x+\xi_{n+1})$ est indépendante de $\mathcal{F}_n$. Ainsi,

$$\mathbb{E}[f(X_{n+1})\,|\,\mathcal{F}_n] = \sum_{x\in E} \mathbb{1}_{\{X_n=x\}}\, \mathbb{E}[f(x+\xi_{n+1})]$$

qui est $\sigma(X_n)$-mesurable. On a donc $\mathbb{E}[f(X_{n+1})\,|\,\mathcal{F}_n] = \mathbb{E}[f(X_{n+1})\,|\,X_n]$, et donc la markoviennité de la marche dans la filtration (naturelle de $X = (X_n)_{n\geqslant 0}$ ou de $X_0, \xi_1, \ldots, \xi_n, \ldots$).

(ii) Supposons les variables aléatoires $(\xi_n)_{n\geqslant 1}$ identiquement distribuées de loi μ sur E. D'après le calcul qui précède

$$\begin{aligned}\mathbb{E}[f(X_{n+1})\,|\,X_n] &= \sum_{x\in E} \mathbb{1}_{\{X_n=x\}}\, \mathbb{E}[f(x+\xi_{n+1})]\\ &= \sum_{x\in E} \mathbb{1}_{\{X_n=x\}} \int_E f(x+\xi)\,\mu(\mathrm{d}\xi)\\ &= \sum_{x\in E} \mathbb{1}_{\{X_n=x\}} \int_E f(\xi)\,\mu(\mathrm{d}\xi - x)\\ &= \int_E f(\xi)\,\mu(\mathrm{d}\xi - X_n) = Pf(X_n).\end{aligned}$$

L'application $P : (x, B) \mapsto \mu(B-x)$ est un noyau de transition sur E (il n'y a aucune difficulté de mesurabilité puisque nous avons supposé E discret), et la formule montre que le processus admet une fonction de transition homogène $(P^n)_{n\geqslant 0}$ donnée par P. La mesure $B \mapsto \mu(B-x)$ est en fait l'image de μ par l'application $y \mapsto y + x$, l'ensemble E n'a donc pas besoin d'être muni d'une soustraction.

Remarque. — Une approche plus générale ($E = \mathbb{R}$ ou $\mathbb{R}^d$) consiste à vérifier la propriété de Markov simple tout d'abord pour $f_\theta(x) = \mathrm{e}^{\mathrm{i}\theta\cdot x}$ pour $\theta \in \mathbb{R}^d$ fixé :

$$\begin{aligned}
\mathbb{E}\big[f_\theta(X_{n+1}) \,\big|\, \mathcal{F}_n\big] &= \mathbb{E}\big[f_\theta(X_n + \xi_{n+1}) \,\big|\, \mathcal{F}_n\big] = \mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\theta\cdot(X_n+\xi_{n+1})} \;\Big|\; \mathcal{F}_n\Big] \\
&= \mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\theta\cdot X_n}\,\mathrm{e}^{\mathrm{i}\theta\cdot\xi_{n+1}} \;\Big|\; \mathcal{F}_n\Big] = \mathrm{e}^{\mathrm{i}\theta\cdot X_n}\,\mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\theta\cdot\xi_{n+1}} \;\Big|\; \mathcal{F}_n\Big] \\
&= \mathrm{e}^{\mathrm{i}\theta\cdot X_n}\,\mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\theta\cdot\xi_{n+1}}\Big] = \mathrm{e}^{\mathrm{i}\theta\cdot X_n}\int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\theta\cdot x}\,\mu_{n+1}(\mathrm{d}x) \\
&= \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\theta\cdot(X_n+x)}\,\mu_{n+1}(\mathrm{d}x) = \int_{\mathbb{R}^d} f_\theta(X_n + x)\,\mu_{n+1}(\mathrm{d}x)
\end{aligned}$$

presque sûrement, qui est $\sigma(X_n)$-mesurable et donc aussi égal à $\mathbb{E}[f_\theta(X_{n+1}) \,|\, X_n]$ (on peut aussi voir que le calcul de $\mathbb{E}[f_\theta(X_{n+1}) \,|\, X_n]$ donne le même résultat), on a

$$\mathbb{E}\big[f_\theta(X_{n+1}) \,\big|\, \mathcal{F}_n\big] = \mathbb{E}\big[f_\theta(X_{n+1}) \,\big|\, X_n\big] = \int_{\mathbb{R}^d} f_\theta(X_n + x)\,\mu_{n+1}(\mathrm{d}x) \qquad \text{presque sûrement.} \quad (*)$$

L'ensemble $\mathcal{H}$ des $f : \mathbb{R}^d \to \mathbb{C}$ boréliennes bornées vérifiant $(*)$ forme un espace vectoriel monotone (modulo la prise des parties réelles et imaginaires), et il est clair qu'il contient les fonctions constantes $\mathbb{C}\mathbb{1}$.

Si $f \in \mathcal{S}(\mathbb{R}^d)$ — l'espace de Schwartz des fonctions à décroissance rapide —, alors $f(x) = \int_{\mathbb{R}^d} g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot x}\,\mathrm{d}\theta$ pour $g \in \mathcal{S}(\mathbb{R}^d)$ (résultat non évident qui est au cœur de la théorie des distributions tempérées de Schwartz). On a alors

$$f(X_{n+1}) = f(X_n + \xi_{n+1}) = \int_{\mathbb{R}^d} g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot(X_n+\xi_{n+1})}\,\mathrm{d}\theta,$$

et on a formellement (*voir* plus loin la justification du « Fubini conditionnel »)

$$\begin{aligned}
\mathbb{E}\big[f(X_{n+1}) \,\big|\, \mathcal{F}_n\big] &= \mathbb{E}\left[\int_{\mathbb{R}^d} g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot(X_n+\xi_{n+1})}\,\mathrm{d}\theta \;\middle|\; \mathcal{F}_n\right] \\
&= \int_{\mathbb{R}^d} \mathbb{E}\Big[g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot(X_n+\xi_{n+1})} \;\Big|\; \mathcal{F}_n\Big]\,\mathrm{d}\theta \qquad \text{(Fubini conditionnel)} \\
&= \int_{\mathbb{R}^d} g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot X_n}\,\mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\theta\cdot\xi_{n+1}} \;\Big|\; \mathcal{F}_n\Big]\,\mathrm{d}\theta \\
&= \int_{\mathbb{R}^d} g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot X_n}\,\mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\theta\cdot\xi_{n+1}}\Big]\,\mathrm{d}\theta \\
&= \int_{\mathbb{R}^d} g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot X_n}\left(\int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\theta\cdot x}\,\mu_{n+1}(\mathrm{d}x)\right)\mathrm{d}\theta \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot(X_n+x)}\,\mathrm{d}\theta\right)\mu_{n+1}(\mathrm{d}x) \qquad \text{(Fubini)} \\
&= \int_{\mathbb{R}^d} f(X_n + x)\,\mu_{n+1}(\mathrm{d}x),
\end{aligned}$$

ce qui montre que f vérifie $(*)$. La classe $\mathcal{C} = \mathbb{C}\mathbb{1} \oplus \mathcal{S}(\mathbb{R}^d)$ est donc contenue dans $\mathcal{H}$; elle contient les constantes et est stable par produit. Donc $\mathcal{H}$ contient l'ensemble des fonctions mesurables bornées mesurables par rapport à $\sigma(\mathcal{C})$ qui est l'ensemble des fonction boréliennes bornées. Donc $\mathcal{H}$ est l'ensemble des fonctions boréliennes bornées. (Attention, comme nous

sommes sur $\mathbb{R}^d$, la tribu de Baire qui est engendrée par les fonctions continues, et par régularisation les fonctions régulières, est égale à la tribu borélienne, alors qu'en général, elle est simplement incluse dans celle-ci. Tout ceci n'est pas si évident…)

Le point délicat dans la démonstration qui précède est le « Fubini conditionnel ». Voici comment le justifier. Soit $A \in \mathcal{F}_n$. Nous donnons comme candidat à être une version de

$$\mathbb{E}\left[\int_{\mathbb{R}^d} g(\theta)\, \mathrm{e}^{\mathrm{i}\theta\cdot(X_n+\xi_{n+1})}\, \mathrm{d}\theta \;\middle|\; \mathcal{F}_n\right]$$

la variable aléatoire

$$Z = \int_{\mathbb{R}^d} g(\theta)\, \mathrm{e}^{\mathrm{i}\theta\cdot X_n} \varphi_{n+1}(\theta)\, \mathrm{d}\theta$$

où

$$\varphi_{n+1}(\theta) = \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\theta\cdot x} \mu_{n+1}(\mathrm{d}x) = \mathbb{E}\big[\mathrm{e}^{\mathrm{i}\theta\cdot\xi_{n+1}}\big]$$

est la fonction caractéristique de la loi μ_{n+1} de l'accroissement ξ_{n+1}. Notons que $(\omega,\theta) \in \Omega\times\mathbb{R}^d \mapsto g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot X_n}\varphi_{n+1}(\theta)$ est mesurable et $\mathbb{P}\otimes\mathrm{d}\theta$ intégrable, ainsi que son produit avec $\omega\in\Omega\mapsto \mathbb{1}_A(\omega)$. On a alors

$$\begin{aligned}
\mathbb{E}[Z\mathbb{1}_A] &= \mathbb{E}\left[\int_{\mathbb{R}^d} g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot X_n}\varphi_{n+1}(\theta)\mathbb{1}_A\,\mathrm{d}\theta\right]\\
&= \int_{\mathbb{R}^d} g(\theta)\varphi_{n+1}(\theta)\,\mathbb{E}\big[\mathrm{e}^{\mathrm{i}\theta\cdot X_n}\mathbb{1}_A\big]\,\mathrm{d}\theta \qquad \text{(Fubini par rapport à } \mathbb{P}\otimes\mathrm{d}\theta)\\
&= \int_{\mathbb{R}^d} g(\theta)\,\mathbb{E}\big[\mathrm{e}^{\mathrm{i}\theta\cdot\xi_{n+1}}\big]\,\mathbb{E}\big[\mathrm{e}^{\mathrm{i}\theta\cdot X_n}\mathbb{1}_A\big]\,\mathrm{d}\theta\\
&= \int_{\mathbb{R}^d} g(\theta)\,\mathbb{E}\big[\mathrm{e}^{\mathrm{i}\theta\cdot\xi_{n+1}}\,\mathrm{e}^{\mathrm{i}\theta\cdot X_n}\mathbb{1}_A\big]\,\mathrm{d}\theta \qquad \text{(indépendance)}\\
&= \int_{\mathbb{R}^d} g(\theta)\,\mathbb{E}\big[\mathrm{e}^{\mathrm{i}\theta\cdot X_{+1}}\mathbb{1}_A\big]\,\mathrm{d}\theta\\
&= \mathbb{E}\left[\int_{\mathbb{R}^d} g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot X_{+1}}\,\mathrm{d}\theta\mathbb{1}_A\right] \qquad \text{(Fubini)}\\
&= \mathbb{E}[f(X_{n+1})\mathbb{1}_A].
\end{aligned}$$

Comme ceci est vrai pour tout $A\in\mathcal{F}_n$, ceci montre l'égalité presque sûre

$$\mathbb{E}[f(X_{n+1})\,|\,\mathcal{F}_n] = Z = \int_{\mathbb{R}^d} g(\theta)\,\mathrm{e}^{\mathrm{i}\theta\cdot X_n}\varphi_{n+1}(\theta)\,\mathrm{d}\theta.$$

La mesurabilité de la fonction en (ω,θ) dans cette dernière intégrale joue un rôle important pour pouvoir utiliser le théorème de Fubini. C'est intimement lié avec la possibilité de choisir mesurablement des versions des espérances conditionnelles (ce qui est non trivial et s'obtient ici par de simple formules analytiques).

Exercice 5. — Soit X un processus markovien admettant une fonction de transition

$$(P_{s,t})_{0\leqslant s\leqslant t}$$

a priori inhomogène. Que dire du processus espace-temps $(X_t,t)_{t\geqslant 0}$?

Correction. — Nous prenons pour espace d'états $\bar{E} = E\times\mathbb{R}_+$ muni de la tribu produit $\bar{\mathcal{E}} = \mathcal{E}\otimes\mathcal{B}(\mathbb{R}_+)$. Le processus espace-temps $\bar{X} = (\bar{X}_t)_{t\geqslant 0} = (X_t,t)_{t\geqslant 0}$ est bien sûr encore

adapté. Pour $f : \bar{E} \to \mathbb{R}$ mesurable bornée, $s \leqslant t$, on a

$$\begin{aligned}\mathbb{E}[f(\bar{X}_t) \,|\, \mathcal{F}_s] = \mathbb{E}[f(X_t, t) \,|\, \mathcal{F}_s] &= \int_E f(y,t) P_{s,t}(X_s, \mathrm{d}y) \\ &= \int_E \Big(\int_{\mathbb{R}_+} f(y,v)\delta_t(\mathrm{d}v) \Big) P_{s,t}(X_s, \mathrm{d}y) \\ &= \int_{\bar{E}} f(y,v) P_{s,t}(X_s, \mathrm{d}y) \otimes \delta_t(\mathrm{d}v) \\ &= \int_{\bar{E}} f(y,v) Q_{t-s}\big((X_s,s), (\mathrm{d}y, \mathrm{d}v)\big) \\ &= \int_{\bar{E}} f(\bar{y}) Q_{t-s}(\bar{X}_s, \mathrm{d}\bar{y})\end{aligned}$$

où plus précisément $Q_h((x,u),(\mathrm{d}y,\mathrm{d}v)) = P_{u,u+h}(x,\mathrm{d}y) \otimes \delta_{u+h}(\mathrm{d}v)$. Ceci montre à la fois que le processus $\bar{X}$ est markovien dans la filtration d'origine et qu'il est homogène admettant $(Q_t)_{t>0}$ pour semi-groupe de transition.

Exercice 6. — Soit P un noyau de probabilité sur un espace d'états $(E, \mathcal{E})$. Une application $f : (E, \mathcal{E}) \to \mathbb{R}$ est dite harmonique (resp. sur, sous-harmonique) si $Pf = f$ (resp. $Pf \leqslant f$, $Pf \geqslant f$).

Soit $X = (X_n)_{n \geqslant 0}$ un processus markovien homogène admettant P pour noyau de transition. Que dire de $f(X) = (f(X_n))_{n \geqslant 0}$ lorsque f est harmonique (resp. sur, sous-harmonique) et vérifie des hypothèses d'intégrabilité convenables ?

Correction. — Si f est harmonique, pour $n \geqslant 0$ et sous réserve d'intégrabilité de $f(X_{n+1})$, $\mathbb{E}[f(X_{n+1}) \,|\, \mathcal{F}_n] = Pf(X_n) = f(X_n)$ presque sûrement. Ainsi $f(X)$, qui est adapté, si $f(X_n)$ est intégrable pour tout $n \geqslant 0$, est alors une martingale.

De même si $Pf \leqslant f$ (elle est au dessus de sa moyenne harmonique), on obtient une sur-martingale, et si $Pf \leqslant f$ (elle est au dessous de sa moyenne harmonique), on obtient une sous-martingale.

Exercice 7. — Nous considérons deux noyaux P et Q sur $\mathbb{R}^d$ homogènes en espace, c'est-à-dire qu'il existe deux mesures de probabilité μ et ν sur $\mathbb{R}^d$ telles que $P(x,\mathrm{d}y) = \mu(\mathrm{d}y - x)$ et $Q(x,\mathrm{d}y) = \nu(\mathrm{d}y - x)$.

(i) Exprimer la mesure de probabilité $(PQ)(0,\mathrm{d}z)$ en fonction de μ et ν.

(ii) Reprendre le calcul pour $(PQ)(x,\mathrm{d}z)$ en fonction de μ et ν avec $x \in \mathbb{R}^d$.

(iii) Constater le lien avec les sommes de variables aléatoires indépendantes à valeurs dans $\mathbb{R}^d$.

Correction. — (i) On a

$$(PQ)(0,\mathrm{d}z) = \int_{\mathbb{R}^d} P(0,\mathrm{d}y) Q(y,\mathrm{d}z) = \int_{\mathbb{R}^d} \mu(\mathrm{d}y)\nu(\mathrm{d}z - y) = \mu * \nu(\mathrm{d}z)$$

le produit de convolution de μ et de ν.

(ii) On a

$$\begin{aligned}(PQ)(x,\mathrm{d}z) &= \int_{\mathbb{R}^d} P(x,\mathrm{d}y) Q(y,\mathrm{d}z) = \int_{\mathbb{R}^d} \mu(\mathrm{d}y - x)\nu(\mathrm{d}z - y) \\ &= \int_{\mathbb{R}^d} \mu(\mathrm{d}y')\nu(\mathrm{d}z - x - y') = \mu * \nu(\mathrm{d}z - x).\end{aligned}$$

(iii) Si U de loi μ, V de loi ν, sont deux variables aléatoires indépendantes, $(PQ)(x,\mathrm{d}z)$ est la loi de $x + U + V$.

Remarque. — L'espace d'états dans l'exercice précédent pourraît être aussi bien $\mathbb{Z}^d$ ou $\mathbb{N}^d$, etc.

Exercice 8 (semi-groupe de Poisson). — Soit $P_t(x, \mathrm{d}y) = \mathrm{e}^{-\lambda t} \sum_{n\geqslant 0} \frac{(\lambda t)^n}{n\,!} \delta_{\{x+n\}}(\mathrm{d}y)$ où $\lambda > 0$ est une constante, $t \geqslant 0$ et $x \in \mathbb{R}$.

(i) Montrer que pour chaque $t \geqslant 0$, P_t est un noyau de probabilité (ou markovien) sur $\mathbb{R}$.

(ii) Montrer que $(P_t)_{t\geqslant 0}$ est un semi-groupe de transition sur $\mathbb{R}$.

Remarque. — Dans cet exercice, nous avons pris pour espace d'états $\mathbb{R}$. Généralement, on considère plutôt $\mathbb{N}$ voire $\mathbb{Z}$.

Correction. — (i) Soit $t \geqslant 0$.

Pour $x \in \mathbb{R}$ fixé, $P_t(x, \mathrm{d}y)$ est la somme d'une série de mesures positives sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. C'est donc une mesure positive et on a :

$$P_t(x, \mathbb{R}) = \mathrm{e}^{-\lambda t} \sum_{n\geqslant 0} \frac{(\lambda t)^n}{n\,!} = 1.$$

Ainsi $P_t(x, \mathrm{d}y)$ est une mesure de probabilité.

Pour $f : \mathbb{R} \to \mathbb{R}$ mesurable bornée fixée, on a

$$P_t f(x) = \mathrm{e}^{-\lambda t} \sum_{n\geqslant 0} \frac{(\lambda t)^n}{n\,!} f(x+n).$$

Comme pour tout $n \geqslant 0$, $x \in \mathbb{R} \mapsto (x+n) \mapsto f(x+n)$ est mesurable par composition, $P_t f : \mathbb{R} \to \mathbb{R}$ est une série convergente de fonctions mesurables, elle est donc mesurable. En particulier, si $f = \mathbb{1}_B$ pour B borélien de $\mathbb{R}$, $x \in \mathbb{R} \mapsto P_t(x, B)$ est mesurable.

Donc P_t est un noyau de probabilité sur $\mathbb{R}$.

(ii) Soient $s, t \geqslant 0$. Par définition de la composition des noyaux, pour $x \in \mathbb{R}$ et $B \in \mathcal{B}(\mathbb{R})$, on a

$$\begin{aligned}
P_s P_t(x, B) &= \int_{\mathbb{R}} P_s(x, \mathrm{d}y) P_t(y, B) \\
&= \int_{\mathbb{R}} \mathrm{e}^{-\lambda s} \sum_{m\geqslant 0} \frac{(\lambda s)^m}{m\,!}\, \delta_{\{x+m\}}(\mathrm{d}y)\, \mathrm{e}^{-\lambda t} \sum_{n\geqslant 0} \frac{(\lambda t)^n}{n\,!}\, \mathbb{1}_B(y+n) \\
&= \int_{\mathbb{R}} \mathrm{e}^{-\lambda(s+t)} \sum_{m\geqslant 0} \sum_{n\geqslant 0} \frac{(\lambda s)^m (\lambda t)^n}{m\,!\; n\,!} \delta_{\{x+m\}}(\mathrm{d}y) \mathbb{1}_B(y+n) \\
&= \mathrm{e}^{-\lambda(s+t)} \sum_{m\geqslant 0} \sum_{n\geqslant 0} \frac{(\lambda s)^m (\lambda t)^n}{m\,!\; n\,!} \int_{\mathbb{R}} \delta_{\{x+m\}}(\mathrm{d}y) \mathbb{1}_B(y+n) \\
&= \mathrm{e}^{-\lambda(s+t)} \sum_{m\geqslant 0} \sum_{n\geqslant 0} \frac{(\lambda s)^m (\lambda t)^n}{m\,!\; n\,!} \mathbb{1}_B(x+m+n)
\end{aligned}$$

tout étant positif, on aura interverti les sens de sommation en utilisant le théorème de Fubini–Tonelli. En changeant d'indices $k = m + n$, $\ell = m$ (ce changement d'indices a été vu en L3 et en M1 premier semestre, des erreurs ne sont donc pas pardonnables), par convergence

commutative, on a

$$\begin{aligned}
P_sP_t(x,B) &= \mathrm{e}^{-\lambda(s+t)} \sum_{k\geqslant 0} \sum_{\ell=0}^{k} \frac{(\lambda s)^\ell (\lambda t)^{k-\ell}}{\ell\,!\,(k-\ell)\,!} \mathbb{1}_B(x+k) \\
&= \mathrm{e}^{-\lambda(s+t)} \sum_{k\geqslant 0} \frac{1}{k\,!} \left(\sum_{\ell=0}^{k} \frac{k\,!}{\ell\,!\,(k-\ell)\,!} (\lambda s)^\ell (\lambda t)^{k-\ell} \right) \mathbb{1}_B(x+k) \\
&= \mathrm{e}^{-\lambda(s+t)} \sum_{k\geqslant 0} \frac{(\lambda s+\lambda t)^k}{k\,!} \mathbb{1}_B(x+k) \\
&= P_{s+t}(x,B).
\end{aligned}$$

La famille de noyaux de transition $(P_t)_{t\geqslant 0}$ vérifie donc la propriété de semi-groupe, c'est un semi-groupe de transition.

Exercice 9 (semi-groupe brownien). — Soit $P_t(x,\mathrm{d}y) = \exp(-(y-x)^2/2t)/\sqrt{2\pi t}\,\mathrm{d}y$ sur $\mathbb{R}$ pour $t>0$, et $P_0(x,\mathrm{d}y) = \delta_{\{x\}}(\mathrm{d}y)$.

(i) Montrer que pour chaque t, P_t est un noyau de probabilité (ou markovien) sur $\mathbb{R}$.

(ii) Montrer que $(P_t)_{t\geqslant 0}$ est un semi-groupe de transition sur $\mathbb{R}$.

Correction. — (i) Soit $t>0$.

Pour $x\in\mathbb{R}$ fixé,

$$B\in\mathcal{B}(\mathbb{R}) \longmapsto P(x,B) = \int_{\mathbb{R}} \mathbb{1}_B(y)\,\mathrm{e}^{-(y-x)^2/2t} \frac{\mathrm{d}y}{\sqrt{2\pi t}}$$

es positive, additive (échange somme intégrale par Fubini-Tonelli), et vérifie

$$P(x,\mathbb{R}) = \int_{\mathbb{R}} \mathrm{e}^{-(y-x)^2/2t} \frac{\mathrm{d}y}{\sqrt{2\pi t}} = \int_{\mathbb{R}} \mathrm{e}^{-u^2/2} \frac{\mathrm{d}u}{\sqrt{2\pi}} = 1,$$

c'est donc bien une mesure de probabilité. Pour $B\in\mathbb{R}$ fixé,

$$x\in\mathbb{R} \longmapsto P(x,B) = \int_{\mathbb{R}} \mathbb{1}_B(y)\,\mathrm{e}^{-(y-x)^2/2t} \frac{\mathrm{d}y}{\sqrt{2\pi t}}$$

est C^∞ (*voir* intégrale à paramètre), elle est notamment borélienne. Ainsi P_t est bien un noyau de transition.

(ii) Sans calcul, on reconnaît que $P_t(x,\mathrm{d}z) = \mu_t(\mathrm{d}z-x)$ avec μ_t la loi normale $\mathcal{N}(0,t)$. Ainsi, $P_sP_t(x,\mathrm{d}z) = \mu_s * \mu_t(\mathrm{d}z-x) = \mu_{s+t}(\mathrm{d}z-x) = P_{s+t}(x,\mathrm{d}z)$ puisque la somme de deux variables aléatoires indépendantes de lois respectives $\mathcal{N}(0,s)$ et $\mathcal{N}(0,t)$ est de loi $\mathcal{N}(0,s+t)$.

Effectuons le calcul direct. Par définition de la composition des noyaux, pour $x\in\mathbb{R}$ et

$B \in \mathcal{B}(\mathbb{R})$, pour $s,t>0$, on a

$$\begin{aligned}
P_sP_t(x,B) &= \int_{\mathbb{R}} P_s(x,\mathrm{d}y)P_t(y,B)\\
&= \int_{\mathbb{R}} P_s(x,\mathrm{d}y)\int_{\mathbb{R}} P_t(y,\mathrm{d}z)\mathbb{1}_B(z)\\
&= \int_{\mathbb{R}} \mathrm{e}^{-(x-y)^2/2s}\left(\int_{\mathbb{R}} \mathbb{1}_B(z)\,\mathrm{e}^{-(y-z)^2/2t}\,\frac{\mathrm{d}z}{\sqrt{2\pi t}}\right)\frac{\mathrm{d}y}{\sqrt{2\pi s}}\\
&= \int_{\mathbb{R}^2} \mathbb{1}_B(z)\,\mathrm{e}^{-(x-y)^2/2s-(y-z)^2/2t}\,\frac{\mathrm{d}y\,\mathrm{d}z}{2\pi\sqrt{st}} \qquad \text{(Fubini–Tonelli)}\\
&= \int_{\mathbb{R}} \mathbb{1}_B(z)\left(\int_{\mathbb{R}} \mathrm{e}^{-(x-y)^2/2s-(y-z)^2/2t}\,\mathrm{d}y\right)\frac{\mathrm{d}z}{2\pi\sqrt{st}} \qquad \text{(Fubini–Tonelli)}\\
&= \int_{\mathbb{R}} \mathbb{1}_B(z)\,\mathrm{e}^{-(x-z)^2/2(s+t)}\left(\int_{\mathbb{R}} \mathrm{e}^{(x-z)^2/2(s+t)-(x-y)^2/2s-(y-z)^2/2t}\,\mathrm{d}y\right)\frac{\mathrm{d}z}{2\pi\sqrt{st}}\\
&= \int_{\mathbb{R}} \mathbb{1}_B(z)\,\mathrm{e}^{-(x-z)^2/2(s+t)}\left(\int_{\mathbb{R}} \mathrm{e}^{E/(2st(s+t))}\,\mathrm{d}y\right)\frac{\mathrm{d}z}{2\pi\sqrt{st}}
\end{aligned}$$

en notant que l'expression à intégrer est mesurable en (y,z) et positive pour appliquer le théorème de Fubini–Tonelli et où l'expression E est

$$\begin{aligned}
E &= st(x-z)^2 - t(s+t)(x-y)^2 - s(s+t)(y-z)^2\\
&= stx^2 - 2stxz + stz^2 - stx^2 + 2stxy - sty^2\\
&\quad - t^2x^2 + 2t^2xy - t^2y^2 - s^2y^2 + 2s^2yz - s^2z^2 - sty^2 + 2styz - stz^2\\
&= -t^2x^2 - (s^2+2st+t^2)y^2 - s^2z^2 + 2(st+t^2)xy - 2stxz + 2(s^2+st)yz\\
&= -t^2x^2 - (s+t)^2y^2 - s^2z^2 + 2t(s+t)xy - 2stxz + 2s(s+t)yz\\
&= -\big(tx-(s+t)y+sz\big)^2 = -\big((s+t)y-tx-sz\big)^2,
\end{aligned}$$

et ainsi

$$P_sP_t(x,B) = \int_{\mathbb{R}} \mathbb{1}_B(z)\,\mathrm{e}^{-(x-z)^2/2(s+t)}\left(\int_{\mathbb{R}} \mathrm{e}^{-(tx-(s+t)y+sz)^2/(2st(s+t))}\,\mathrm{d}y\right)\frac{\mathrm{d}z}{2\pi\sqrt{st}}\,.$$

En effectuant le changement de variable

$$u = \frac{(s+t)y-tx-sz}{\sqrt{st(s+t)}} = \sqrt{\frac{s+t}{st}}\,y - \frac{tx+sz}{\sqrt{st(s+t)}}\,, \qquad \mathrm{d}u = \sqrt{\frac{s+t}{st}}\,\mathrm{d}y,$$

qui, pour $z\in\mathbb{R}$ fixé (ainsi que x), décrit $]-\infty,+\infty[$ lorsque y décrit $]-\infty,+\infty[$, on a

$$\begin{aligned}
P_sP_t(x,B) &= \int_{\mathbb{R}} \mathbb{1}_B(z)\,\mathrm{e}^{-(x-z)^2/2(s+t)}\left(\int_{\mathbb{R}} \mathrm{e}^{-u^2/2}\,\mathrm{d}u\times\sqrt{\frac{st}{s+t}}\right)\frac{\mathrm{d}z}{2\pi\sqrt{st}}\\
&= \int_{\mathbb{R}} \mathbb{1}_B(z)\,\mathrm{e}^{-(x-z)^2/2(s+t)}\left(\int_{\mathbb{R}} \mathrm{e}^{-u^2/2}\,\mathrm{d}u\right)\frac{\mathrm{d}z}{2\pi\sqrt{s+t}}\\
&= \int_{\mathbb{R}} \mathbb{1}_B(z)\,\mathrm{e}^{-(x-z)^2/2(s+t)}\sqrt{2\pi}\,\frac{\mathrm{d}z}{2\pi\sqrt{s+t}}\\
&= \int_{\mathbb{R}} \mathbb{1}_B(z)\,\mathrm{e}^{-(x-z)^2/2(s+t)}\,\frac{\mathrm{d}z}{\sqrt{2\pi(s+t)}}\\
&= P_{s+t}(x,B).
\end{aligned}$$

La famille de noyaux de transition $(P_t)_{t\geqslant 0}$ vérifie donc la propriété de semi-groupe, c'est un semi-groupe de transition.

FEUILLE D'EXERCICES N° 4. — CHAÎNES DE MARKOV À ESPACE D'ÉTATS DISCRET

EXERCICE 1. — (i) Soient A, B et $C \in \mathcal{A}$ tels que $\mathbb{P}(B \cap C) > 0$. Montrer que A et B sont indépendants conditionnellement à C si et seulement si

$$\mathbb{P}(A \,|\, B \cap C) = \mathbb{P}(A \,|\, C).$$

On pourra noter au passage qu'ici on a nécessairement $\mathbb{P}(C) > 0$.

(ii) Soient E, F et G des ensembles au plus dénombrables munis de leurs tribus discrètes, et $X : (\Omega, \mathcal{A}, \mathbb{P}) \to E$, $Y : (\Omega, \mathcal{A}, \mathbb{P}) \to F$ et $Z : (\Omega, \mathcal{A}, \mathbb{P}) \to G$ trois variables aléatoires (discrètes, donc).

Montrer que X et Y sont indépendantes conditionnellement à Z si et seulement si, pour tout $x \in E$, $y \in F$, $z \in G$ tel que $\mathbb{P}\{Z = z\} > 0$,

$$\mathbb{P}\{X = x, Y = y \,|\, Z = z\} = \mathbb{P}\{X = x \,|\, Z = z\} \times \mathbb{P}\{Y = y \,|\, Z = z\}.$$

Correction. — Non disponible.

EXERCICE 2. — Soit $(X_n)_{n \geqslant 0}$ une suite de variables aléatoires à valeurs dans E. Pour tout $n \geqslant 0$, soient $\mathcal{F}_n = \sigma\{X_0, \ldots, X_n\}$ et $\mathcal{G}_n = \sigma\{X_n, X_{n+1}, \ldots\}$ les tribus respectivement du passé et du futur à l'instant n.

Montrer que $(X_n)_{n \geqslant 0}$ est une chaîne de Markov si et seulement si pour tout $n \geqslant 0$, $x \in E$, les tribus $\mathcal{F}_n$ et $\mathcal{G}_n$ sont indépendantes conditionnellement à $\{X_n = x\}$.

Correction. — Non disponible.

EXERCICE 3 (ALGORITHMES MARKOVIENS). — Dans cet exercice, nous supposerons l'espace d'états E fini ou dénombrable muni de sa tribu discrète. Cette restriction peut être largement levée avec du travail.

(i) Soient $U = (U_1, \ldots, U_n)$ un n-uplet de variables aléatoires indépendantes à valeurs dans un espace mesurable $(F, \mathcal{F})$, et $f_1, \ldots, f_n : F \times E \to E$ des applications mesurables. Soit X_0 une variable aléatoire à valeurs dans E indépendante de U. En posant

$$X_k = f_k(U_k, X_{k-1}), \qquad 1 \leqslant k \leqslant n,$$

alors $(X_k)_{0 \leqslant k \leqslant n}$ est une chaîne de Markov finie.

(ii) Réciproquement, on peut réaliser toute chaîne de Markov de cette manière : étant donnée la loi d'une chaîne de Markov $(Y_0, \ldots, Y_n)$, c'est-à-dire la donnée de la mesure de probabilité initiale λ, et des matrices de transition $P_{0,1}, \ldots, P_{n-1,n}$, il est possible de construire de la manière précédente une chaîne de Markov ayant même loi.

Correction. — (i) Pour tout $k \in \{0, \ldots, n-1\}$, X_k est une fonction de $(X_0, U_1, \ldots, U_k)$ et

est donc indépendante de U_{k+1}. On a alors

$$\begin{aligned}
\mathbb{P}\{X_{k+1}=x_{k+1}\,|\,X_0=x_0,\dots,X_k=x_k\} &= \mathbb{P}\{f_k(U_{k+1},X_k)=x_{k+1}\,|\,X_0=x_0,\dots,X_k=x_k\}\\
&= \mathbb{P}\{f_k(U_{k+1},x_k)=x_{k+1}\,|\,X_0=x_0,\dots,X_k=x_k\}\\
(\text{indépendance de } U_{k+1} \text{ et } (X_0,\dots,X_k)) &= \mathbb{P}\{f_k(U_{k+1},x_k)=x_{k+1}\}\\
(\text{indépendance de } U_{k+1} \text{ et } X_k) &= \mathbb{P}\{f_k(U_{k+1},x_k)=x_{k+1}\,|\,X_k=x_k\}\\
&= \mathbb{P}\{f_k(U_{k+1},X_k)=x_{k+1}\,|\,X_k=x_k\}\\
&= \mathbb{P}\{X_{k+1}=x_{k+1}\,|\,X_k=x_k\}.
\end{aligned}$$

D'où la propriété de Markov.

(ii) Réciproquement, on se donne X_0 et $(U_1,\dots,U_n)$ indépendantes telles que X_0 est à valeurs dans E et de loi λ et chaque U_k est à valeurs dans $F=E^E$ muni de $\mathcal{F}=\mathcal{P}(E)^{\otimes E}$ sa tribu produit (tout point $u\in F$ est une application de E dans E) et dont la loi satisfait la condition

$$\mathbb{P}\{U_k(x)=y\}=P_{k-1,k}(x,y),\qquad x,y\in E$$

(il suffit pour cela que la loi de U_k soit la loi produit $\bigotimes_{x\in E}P_{k-1,k}(x,.)$). On définit ensuite $f:F\times E\to E$ par $f(u,x)=u(x)$ qui est bien entendu mesurable pour la tribu produit. La suite $(X_0,\ X_k=f(U_k,X_{k-1}),1\leqslant k\leqslant n)$ est markovienne d'après le point (i), et, pour x, $y\in E$ et $1\leqslant k\leqslant n$, on a

$$\begin{aligned}
\mathbb{P}\{X_k=y\,|\,X_{k-1}=x\} &= \mathbb{P}\{f(U_k,X_{k-1})=y\,|\,X_{k-1}=x\}\\
&= \mathbb{P}\{f(U_k,x)=y\,|\,X_{k-1}=x\}=\mathbb{P}\{f(U_k,x)=y\}=\mathbb{P}\{U_k(x)=y\}=P_{k-1,k}(x,y)
\end{aligned}$$

où on se sera servi du fait que U_k et X_{k-1} sont indépendantes. □

Remarques. — *a*) Dans la construction précédente nous n'avons pas explicité sur quel espace probabilisé vivaient X_0 et $(U_1,\dots,U_n)$. Il suffit de prendre $\Omega=E\otimes F\otimes\cdots\otimes F$ muni de sa tribu produit et de la mesure de probabilité

$$\lambda\otimes\left(\bigotimes_{x\in E}P_{0,1}(x,.)\right)\otimes\cdots\otimes\left(\bigotimes_{x\in E}P_{n-1,n}(x,.)\right),$$

les variables aléatoires étant alors les projections canoniques correspondant à ce produit.

b) Lorsque E est fini, E^E est lui aussi fini et sa tribu produit est sa tribu discrète. En revanche ce n'est plus le cas si E est infini dénombrable, E^E a alors la puissance du continu (le même cardinal que $\mathbb{R}$), mais cela reste suffisamment raisonnable pour même dans ce cas effectuer des simulations réalistes sur ordinateur.

c) Le point (ii) montre qu'avec la donnée de λ, $P_{0,1},\dots,P_{n-1,n}$, on peut construire une chaîne de Markov ayant cette loi initiale et ces matrices de transition. C'est donc un *théorème d'existence.*

Exercice 4. — Considérons la chaîne de Markov à 2 états de matrice de transition

$$P=\begin{pmatrix}1-\alpha & \alpha\\ \beta & 1-\beta\end{pmatrix}$$

avec α, $\beta\in[0,1]$.

(i) Discuter en fonction de α et de β des classes communicantes et de leur fermeture.

(ii) Calculer les coefficients $P^n(x,y)$.

(iii) Le critère de récurrence $\sum_{n\geqslant 0}P^n(x,x)$ est-il cohérent avec l'analyse faite en (i) ?

Correction. — Non disponible.

Exercice 5. — On considère une chaîne de Markov à 3 états $\{1,2,3\}$ de matrice de transition

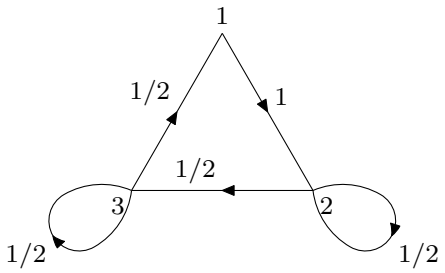
$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix}.$$

Trouver une formule générale pour la suite $(P^n(1,1))_{n\geqslant 0}$. (*Indication :* on pourra se servir des valeurs propres de P.)

Correction. — Le calcul bête et méchant :

```
-->P=[0,1,0;0,0.5,0.5;0.5,0,0.5]; det(%s*eye(3,3)-P)
```

donne $\chi_P(x) = x^3 - x^2 + x/4 - 1/4 = (x^2+1/4)(x-1) = (x-\mathrm{i}/2)(x+\mathrm{i}/2)(x-1)$. On a donc $P^n(1,1) = a\times 1^n + b\times(\mathrm{i}/2)^n + c\times(-\mathrm{i}/2)^n$. Comme $P^0(1,1)=1$, $P^1(1,1)=0$ et $P^2(1,1)=0$ comme on le voit sur le graphe



On résout

$$\begin{cases} a+b+c=1 \\ a+b\mathrm{i}/2-c\mathrm{i}/2=0 \\ a-b/4-c/4=0 \end{cases} \quad \begin{cases} 5a=1 \\ a+b\mathrm{i}/2-c\mathrm{i}/2=0 \\ a-b/4-c/4=0 \end{cases} \quad \begin{cases} a=1/5 \\ b-c=2\mathrm{i}/5 \\ b+c=4/5 \end{cases} \quad \begin{cases} a=1/5 \\ b=(2+\mathrm{i})/5 \\ c=(2-\mathrm{i})/5 \end{cases}$$

D'où

$$P^n(1,1) = \frac{1}{5} + \frac{2+\mathrm{i}}{5}\Big(\frac{\mathrm{i}}{2}\Big)^n + \frac{2-\mathrm{i}}{5}\Big(-\frac{\mathrm{i}}{2}\Big)^n$$

ce qu'on pourrait écrire sous forme réelle...

Exercice 6. — Soient $(X_n)_{n\geqslant 0}$ est une chaîne de Markov de matrice de transition P (et de loi initiale λ) et k entier strictement positif. Montrer que si on pose $Y_n = X_{kn}$, pour tout $n \geqslant 0$, alors la suite de variables aléatoires $(Y_n)_{n\geqslant 0}$ est une chaîne de Markov de matrice de transition P^k (et de loi initiale λ).

Correction. — Non disponible.

Exercice 7. — On suppose que la matrice de transition P de X n'a pas d'état absorbant. On définit

$$S_0 = 0, \qquad S_{m+1} = \inf\big\{n \geqslant S_m : X_n \neq X_{S_m}\big\}, \quad m \geqslant 0.$$

Montrer que $(S_m)_{m\geqslant 0}$ est une suite de temps d'arrêt, finis.

En posant $Z_m = X_{S_m}$, on définit alors une nouvelle chaîne de Markov : c'est la chaîne $(X_n)_{n\geqslant 0}$ observée uniquement lorsqu'elle « bouge ». Montrer que sa matrice de transition $\bar{P}$ vérifie

$$\bar{P}(x,x) = 0 \quad \text{pour tout } x \in E, \qquad \text{et} \qquad \bar{P}(x,y) = \frac{P(x,y)}{\sum_{z\neq x} P(x,z)}$$

pour tout $x, y \in E$ tels que $x \neq y$.

Exercice 8 (Marches aléatoires simples dans $\mathbb{Z}$). — On considère pour espace d'états $E = \mathbb{Z}$ et $p \in [0,1]$ un réel. En posant $q = 1 - p$, on définit une matrice de transition P par

$$P(x,y) = \begin{cases} p & \text{si } y = x+1 \\ q & \text{si } y = x-1 \\ 0 & \text{sinon.} \end{cases}$$

Montrer que

$$P^n(x,y) = \begin{cases} \binom{n}{(n+y-x)/2} p^{(n+y-x)/2} q^{(n-y-x)/2} & \text{si } n+y-x \text{ est pair} \\ 0 & \text{sinon,} \end{cases}$$

où $\binom{n}{k} = \frac{n\,!}{k\,!\,(n-k)\,!}$ pour $k \in \{0, \dots, n\}$ et 0 sinon.

Correction. — Non disponible.

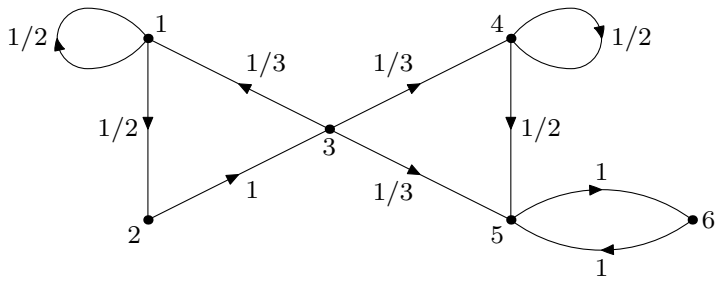
Exercice 9. — Dans le cas d'une chaîne de Markov à deux états, discuter en fonction de α et β des relations entre les deux points, des classes communicantes, des classes fermées.

Correction. — Non disponible.

Exercice 10. — Identifier les classes communicantes, en précisant lesquelles sont fermées, associées respectivement aux matrices de transition

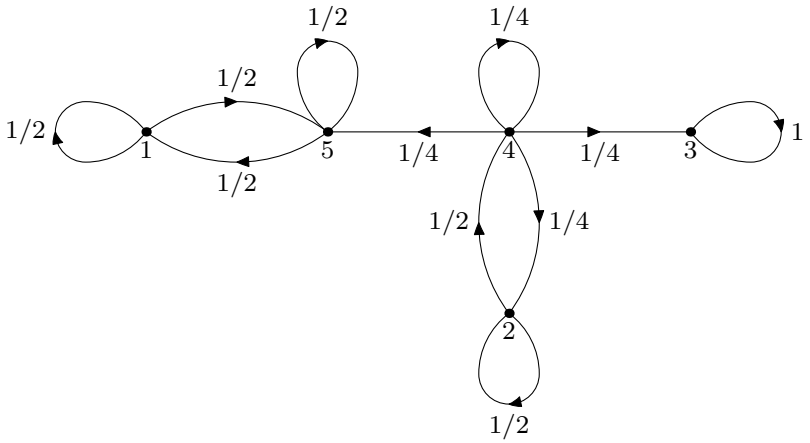
$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \qquad Q = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix}.$$

Correction. — L'espace d'états est $E = \{1,2,3,4,5,6\}$. La matrice de transition P correspond au diagramme de transition suivant :



Les classes communicantes sont $\{1,2,3\}$, $\{4\}$ et $\{5,6\}$ — la chaîne n'est pas irréductible —, et seule cette dernière classe est fermée.

Le diagramme de transition associé à la matrice de transition Q est



Les classes communicantes sont $\{1,5\}$, $\{3\}$, $\{2,4\}$, et seules les deux premières sont fermées. L'état 3 est absorbant.

Exercice 11. — Montrer que toute matrice de transition sur un espace d'état fini a au moins une classe fermée. Trouver un exemple de matrice de transition sans classe fermée.

Correction. — Soient E un espace d'états (non vide) et P une matrice de transition tels qu'aucune classe communicante est fermée. Pour $x \in E$, nous notons $\dot{x} \subset E$ sa classe communicante. Soit $x_1 \in E$. Puisque $\dot{x}_1$ n'est pas fermée, soit $x_2 \in E \backslash \dot{x}_1$ tel que $x_1 \rightsquigarrow x_2$. Par récurrence, pour tout $n \geqslant 1$, nous pouvons construire une suite $x_1, \dots, x_n$ dans E telle que pour tout $1 \leqslant i < n$, $x_{i+1} \in E \backslash \dot{x}_i$ et $x_i \rightsquigarrow x_{i+1}$. Une telle suite est nécessairement injective puisque si pour $1 \leqslant i < j \leqslant n$, on a $x_i = x_j$, comme par construction $x_i \rightsquigarrow x_{i+1}$ et $x_{i+1} \rightsquigarrow x_j = x_i$ alors $x_{i+1} \in \dot{x}_i$. Ainsi, pour tout $n \geqslant 1$, il existe une injection de $\{1, \dots, n\}$ dans E. L'ensemble E est donc infini.

Nous aurions pu améliorer un peu ce résultat en montrant que si toute les classes sont ouvertes, alors le nombre de classes communicantes est infini. Par contraposée, si le nombre de classes est fini, alors il existe au moins une classe fermée.

Pour exemple de matrice de transition sans classe fermée, prenons $E = \mathbb{Z}$ et pour chaîne la marche déterministe consistant à toujours faire un pas à droite. La matrice de transition $P = (P(x,y))_{x,y \in \mathbb{Z}}$ est définie par $P(x, x+1) = 1$, $x \in \mathbb{Z}$, les autres coefficients étant évidemment nuls. Si $x \in \mathbb{Z}$, alors $x \rightsquigarrow y$ pour tout $y \geqslant x$ mais $x \not\rightsquigarrow y$ si $y < x$. Les classes communicantes sont réduites au singletons $\{x\}$, $x \in \mathbb{Z}$, qui sont toutes ouvertes.

Université de Poitiers
Département de Mathématiques

FEUILLE D'EXERCICES N° 5. — CHAÎNES DE MARKOV À ESPACE D'ÉTATS DISCRET (SUITE)

Exercice 1. — Soient E un espace d'états au plus dénombrable et P une matrice de transition.

(i) Montrer que si P est bistochastique, alors les mesures uniformes sont invariantes.

(ii) En déduire que si P est symétrique ($P(x,y) = P(y,x)$ pour tous $x, y \in E$), alors les mesures uniformes sont invariantes.

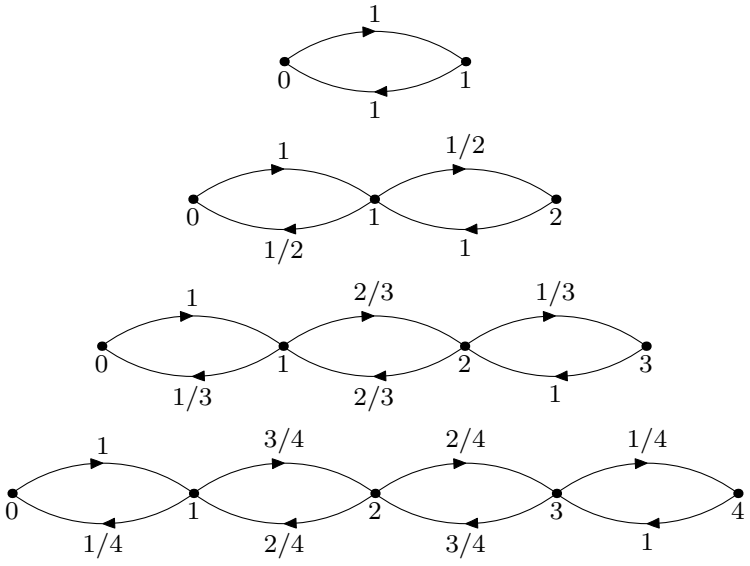
Correction. — Non disponible.

Exercice 2 (l'urne d'Ehrenfest, résumé et suite). — Nous considérons deux urnes A et B et N boules numérotées de 1 à N réparties dans les deux urnes. On tire uniformément un nombre entre 1 et N. La boule portant ce numéro est alors déplacée dans l'autre urne (si elle est dans A, elle va dans B, si elle est dans B, elle va dans A). Et on continue...

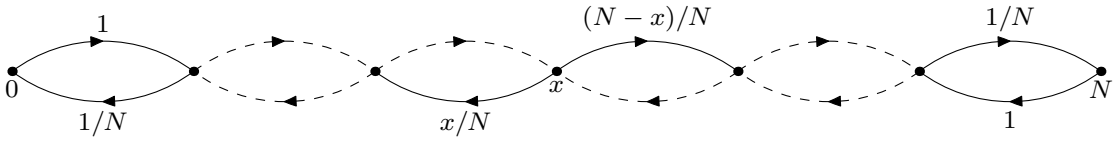
Soit X_n le nombre de boules dans l'urne A à l'étape n, qui est une variable aléatoire à valeurs dans $E = \{0, 1, \dots, N\}$. En admettant que $(X_n)_{n\geqslant 0}$ est bien une chaîne de Markov homogène, sa matrice de transition s'écrit alors

$$P = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ \frac{1}{N} & 0 & \frac{N-1}{N} & \ddots & & & \vdots \\ 0 & \frac{2}{N} & 0 & \frac{N-2}{N} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \frac{N-2}{N} & 0 & \frac{2}{N} & 0 \\ \vdots & & & \ddots & \frac{N-1}{N} & 0 & \frac{1}{N} \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Voici quelques diagrammes de transition pour $N = 1, \dots, 4$:



et plus généralement



(i) Classification des états : les diagrammes montrent clairement que la chaîne est irréductible. Il est aussi facile d'expliciter un chemin joignant deux points x et y quelconques de l'espace d'états $E = \{0, \dots, N\}$ et ainsi de trouver une probabilité $P^n(x, y)$ qui soit strictement positive. L'espace d'états étant fini, tous les états sont récurrents.

(ii) La matrice de transition P est réversible par rapport à la mesure de probabilité π définie par

$$\pi\{x\} = \binom{N}{x} \times \frac{1}{2^N}\,, \qquad x \in \{0, 1, \dots, N\}.$$

En effet, si $0 \leqslant x < N$, on a

$$\pi\{x\} \times P(x, x+1) = \binom{N}{x} \times \frac{1}{2^N} \times \frac{N-x}{N} = \binom{N-1}{x} \times \frac{1}{2^N}$$

et

$$\pi\{x+1\} \times P(x+1, x) = \binom{N}{x+1} \times \frac{1}{2^N} \times \frac{x+1}{N} = \binom{N-1}{x} \times \frac{1}{2^N},$$

et toutes les autres relations envisageables donnent l'égalité de 0 avec lui-même car $P(x, y) = 0$ si $x = y$ ou $|x - y| \geqslant 2$. Comme P est réversible par rapport à la mesure de probabilité π, cette mesure est aussi invariante. On en déduit que π, qui est la loi binomiale $\mathcal{B}(N, 1/2)$, est l'unique mesure de probabilité invariante.

(iii) La chaîne est-t-elle apériodique ? Que se passe-t-il simplement qui prévient d'une convergence vers l'équilibre ?

(iv) On s'intéresse à la matrice de transition P^2. Puisque π est invariante par P, elle est aussi invariante par P^2. Montrer que P^2 est réversible par rapport à π.

Soit E_p l'ensemble des états pairs, E_i l'ensemble des états impairs. Posons $Q_p = P^2_{|E_p}$ et $Q_i = P^2_{|E_i}$. Le symbole $*$ désignera l'une ou l'autre des lettres $\{p, i\}$.

(v) Montrer que Q_* est irréductible et récurrente, que $\pi_* = \pi_{|E_*}/\pi(E_*)$ est une mesure de probabilité par rapport à laquelle Q_* est réversible donc invariante de Q_*.

(vi) Montrer que Q_* est apérodique. En déduire la convergence vers l'équilibre : pour toute mesure de probabilité λ_* portée par E_*, on a $\lambda_* Q_*^n \to \pi_*$.

(vii) En déduire que pour toute mesure de probabilité λ sur E, $\lambda P^{2n} \to \lambda(E_i)\pi_i + \lambda(E_p)\pi_p$ et $\lambda P^{2n+1} \to \lambda(E_i)\pi_p + \lambda(E_p)\pi_i$ et qu'ainsi $\lambda(P^n + P^{n+1})/2$ converge quand n tend vers l'infini (la limite est presque π mais pas tout à fait).

(viii) Le théorème ergodique s'applique-t-il ? Qu'affirme-t-il ?

Correction. — Non disponible.

Exercice 3. — Considérons un graphe connexe fini $(E, \mathcal{R})$ (la relation $\mathcal{R}$ est nulle part réflexive). Soit P définie par

$$P(x, y) = \begin{cases} \dfrac{1}{v\{x\} + 1} & \text{si } x \,\mathcal{R}\, y \text{ ou } x = y, \\ 0 & \text{sinon.} \end{cases}$$

Montrer que P est une matrice stochastique irréductible, récurrente et apériodique. Déterminer sa mesure de probabilité invariante.

Correction. — Puisque le graphe est connexe, la matrice P est irréductible : si $x, y \in E$, alors il existe $x_0 = x, x_1, \ldots, x_{n-1}, x_n = y$ tels que $x \mathcal{R} x_1, \ldots, x_{n-1} \mathcal{R} y$. Ainsi, $P^n(x,y) \geqslant P(x,x_1) \times \cdots \times P(x_{n-1}, y) = \frac{1}{v\{x\}+1} \times \ldots \times \frac{1}{v\{x_{n-1}\}+1} > 0$. Comme E est fini, P est récurrente. Que P soit apériodique est une conséquence immédiate du fait que $P(x,x) > 0$ pour tout $x \in E$, ce qui, on le rappelle, pourrait ne pas être garanti si nous n'avions pas convenu que chaque sommet était, finalement, joint à lui-même.

Le cardinal de E étant fini, la récurrence est positive et il existe une unique mesure de probabilité invariante. La mesure positive μ donnée par $\mu\{x\} = v\{x\} + 1$ est clairement telle que P est réversible par rapport à μ. Convenablement normalisée, elle donne la mesure invariante :

$$\pi\{x\} = \frac{v\{x\}+1}{\sum_{y\in E}(v\{y\}+1)} = \frac{v\{x\}+1}{\sum_{y\in E} v\{y\} + \operatorname{Card} E}\,.$$

Exercice 4. — Une particule se déplace sur les 8 sommets d'un cube de la manière suivante : à chaque pas la particule a la même probabilité d'aller sur chacun des trois sommets adjacents, indépendamment de ce qui a pu se passer auparavant. Soit i le sommet initialement occupé par la particule, o le sommet opposé. Calculer chacune des quantités suivantes :

(i) le nombre moyen de pas avant que la particule ne revienne en i ;

(ii) le nombre moyen de visites de o avant son premier retour en i ;

(iii) le nombre moyen de pas avant la première visite de o.

Correction. — (i) *Espérance du temps de retour en i.* La chaîne est irréductible et a pour mesure invariante π la mesure uniforme ($\pi\{s\} = 1/8$ pour tout sommet s) (*voir* la discussion du cours sur les marches aléatoires sur un graphe fini). Alors l'espérance du temps de retour en i est

$$m_{\mathrm{r}}(i) = \mathbb{E}^i[T_{\mathrm{r}}^i] = 1/\pi\{i\} = 8$$

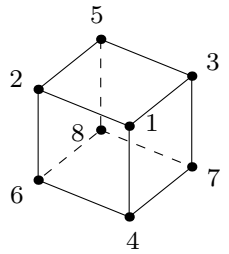
(*voir* le cours, chap. 3, §5. Mesures invariantes, récurrence positive).

(ii) *Nombre moyen de visites de o avant de revenir en i.* Ce que nous cherchons est

$$\gamma^i\{o\} = \mathbb{E}^i\Bigg[\sum_{n=0}^{T_{\mathrm{r}}^i-1} \mathbb{1}_{\{X_n=o\}}\Bigg].$$

On a par irreductibilité et récurrence, $\gamma^i P = \gamma^i$, avec $\gamma^i\{i\} = 1$, donc on a $\gamma^i\{o\} = 1$.

(iii) *Le nombre moyen de pas avant la première visite de o.* Nous cherchons $\mathbb{E}^i[T_{\mathrm{r}}^o]$. Pour fixer les notations, prenons pour point initial $i = 1$ et pour point opposé $o = 8$.



Plus généralement, posons $m_{\mathrm{r}}(j,k) = \mathbb{E}^j[T_{\mathrm{r}}^k]$. En faisant 1 pas et en appliquant la propriété de Markov simple,

$$m_{\mathrm{r}}(j,k) = 1 + \sum_{\ell \neq k} P(j,\ell) m_{\mathrm{r}}(\ell,k).$$

On pose $m_{\mathrm{r}}(j) = m_{\mathrm{r}}(j, 8)$. On sait déjà que $m_{\mathrm{r}}(8) = 8$ et par symétrie $m_{\mathrm{r}}(2) = m_{\mathrm{r}}(3) = m_{\mathrm{r}}(4)$ et $m_{\mathrm{r}}(5) = m_{\mathrm{r}}(6) = m_{\mathrm{r}}(7)$. Le système s'écrit

$$\begin{cases} m_{\mathrm{r}}(1) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(2) + m_{\mathrm{r}}(3) + m_{\mathrm{r}}(4)) = 1 + m_{\mathrm{r}}(2) \\ m_{\mathrm{r}}(2) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(1) + m_{\mathrm{r}}(5) + m_{\mathrm{r}}(6)) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(1) + 2m_{\mathrm{r}}(5)) \\ m_{\mathrm{r}}(3) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(1) + m_{\mathrm{r}}(5) + m_{\mathrm{r}}(7)) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(1) + 2m_{\mathrm{r}}(5)) \\ m_{\mathrm{r}}(4) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(1) + m_{\mathrm{r}}(6) + m_{\mathrm{r}}(7)) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(1) + 2m_{\mathrm{r}}(5)) \\ m_{\mathrm{r}}(5) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(2) + m_{\mathrm{r}}(3)) = 1 + \dfrac{2}{3}\, m_{\mathrm{r}}(2) \\ m_{\mathrm{r}}(6) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(2) + m_{\mathrm{r}}(4)) = 1 + \dfrac{2}{3}\, m_{\mathrm{r}}(2) \\ m_{\mathrm{r}}(7) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(3) + m_{\mathrm{r}}(4)) = 1 + \dfrac{2}{3}\, m_{\mathrm{r}}(2) \\ m_{\mathrm{r}}(8) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(5) + m_{\mathrm{r}}(6) + m_{\mathrm{r}}(7)) \quad (= 8 \quad \text{pour vérification uniquement}) \end{cases}$$

Il ne reste donc que

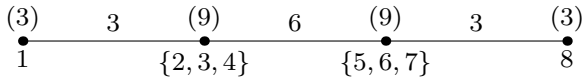
$$\begin{cases} m_{\mathrm{r}}(1) = 1 + m_{\mathrm{r}}(2) \\ m_{\mathrm{r}}(2) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(1) + 2m_{\mathrm{r}}(5)) \\ m_{\mathrm{r}}(5) = 1 + \dfrac{2}{3}\, m_{\mathrm{r}}(2) \end{cases}$$

qui a pour unique solution

$$\begin{cases} m_{\mathrm{r}}(1) = 10 \\ m_{\mathrm{r}}(2) = 9 \\ m_{\mathrm{r}}(5) = 7 \end{cases}$$

D'où la solution.

Autre méthode. — On replie les sommets $\{2, 3, 4\}$ et $\{5, 6, 7\}$ sur eux-mêmes pour obtenir le graphe pondéré



La somme des valences est 24. La mesure invariante est $(3/24, 9/24, 9/24, 3/24) = \frac{1}{8}(1, 3, 3, 1)$. On retrouve $\mathbb{E}^1[T^1_{\mathrm{r}}] = \mathbb{E}^8[T^8_{\mathrm{r}}] = 8$. On réécrit le système associé

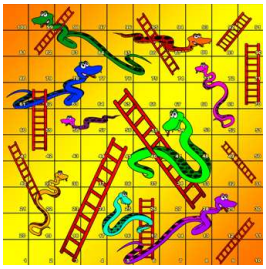
$$m_{\mathrm{r}}(j, k) = 1 + \sum_{\ell \neq k} P(j, \ell) m_{\mathrm{r}}(\ell, k),$$

soit

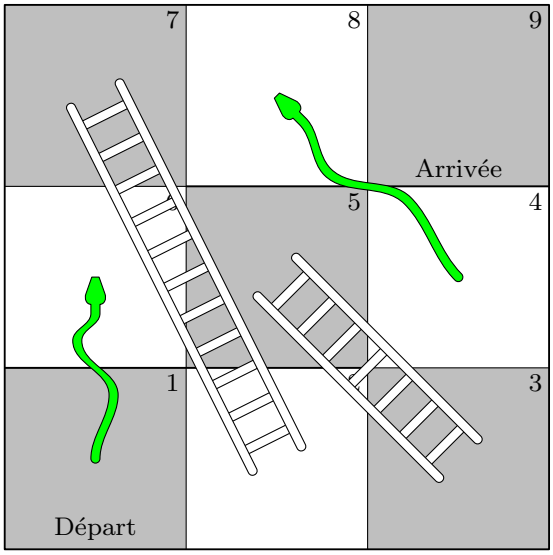
$$\begin{cases} m_{\mathrm{r}}(1) = 1 + m_{\mathrm{r}}(\{234\}) \\ m_{\mathrm{r}}(\{234\}) = 1 + \dfrac{1}{3}(m_{\mathrm{r}}(1) + 2m_{\mathrm{r}}(\{567\})) \\ m_{\mathrm{r}}(\{567\}) = 1 + \dfrac{1}{3}(0 + 2m_{\mathrm{r}}(\{234\})) \qquad (\text{le } 0 \text{ remplace } m_{\mathrm{r}}(8) \text{ qui est interdit}) \\ m_{\mathrm{r}}(8) = 1 + m_{\mathrm{r}}(\{5, 6, 7\}) \quad (= 8 \quad \text{pour vérification}) \end{cases}$$

On retrouve le même système d'équations que précédemment, donc la même solution.

Exercice 5 (*Snakes & Ladders*). — Le jeu « Snakes & Ladders » (serpents et échelles) est un jeu anglo-saxon destiné aux jeunes enfants. On dispose d'un plateau carré divisé en $n \times n$ cases numérotées où sont dessinés des serpents et des échelles joignant certaines cases. Partant de la case 1, on lance un dé et avance d'autant de cases dans l'ordre de leur numérotation. Si la case atteinte correspond à une tête de serpent, on saute immédiatement à la case correspondant à l'extrémité de sa queue. Si la case correspond à un pied d'échelle, on saute immédiatement à la case correspondant au haut de l'échelle. Si la case est vierge, on ne fait rien. Ceci peut se jouer à plusieurs, le gagnant étant celui qui aura atteint ou dépassé la dernière case en effectuant le moins de tours de jeu. Des variantes plus ou moins éducatives existent.



Nous considérons une version simplifiée de « Snakes & Ladders ». Le plateau de jeu comporte 9 cases et est représenté ci-dessous



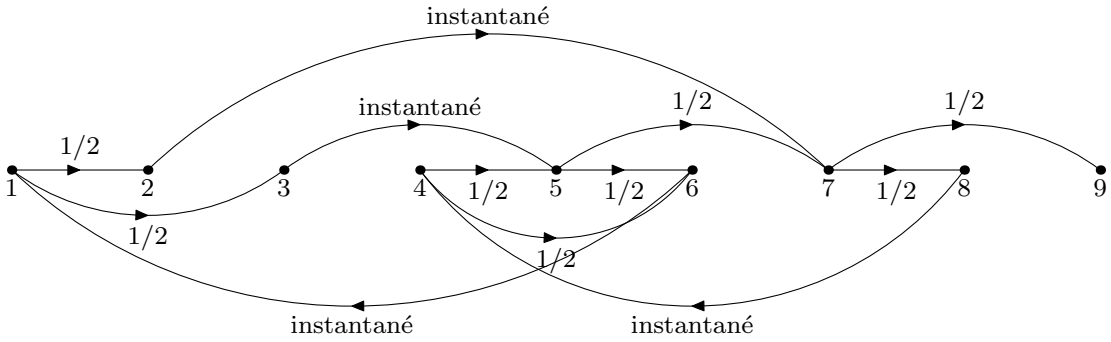
La progression se fait en lançant une pièce de monnaie : en obtenant pile, on avance d'une case ; en obtenant face, on avance de deux cases.

(i) En moyenne, combien faut-il de tours pour terminer la partie ?

(ii) Quelle est la probabilité qu'un joueur qui a atteint la case du milieu termine le jeu sans retourner à la case de départ ?

Correction. — Le déroulement du jeu s'apparente à une chaîne de Markov dès lors qu'on

modélise correctement les transitions instantanées.



(i) Commençons par suivre ce premier diagramme. Soient $m_{\mathrm{a}}(x) = m_{\mathrm{a}}^9(x) = \mathbb{E}^x[T^9]$, $i = 1, \dots, 9$. On a clairement $m_{\mathrm{a}}(9) = 0$. Ensuite, lorsque x est différent de 9 ou d'une case de transition immédiate, la propriété de Markov simple montre que

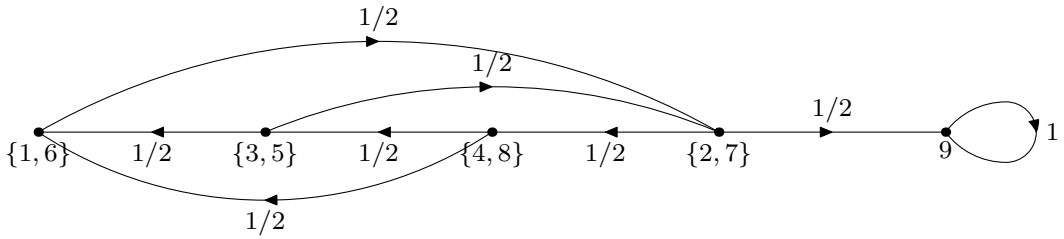
$$m_{\mathrm{a}}(x) = 1 + \sum_{y=1}^{9} P(x,y) m_{\mathrm{a}}(y).$$

Enfin, si x correspond à une case de transition immédiate $m_{\mathrm{a}}(x) = m_{\mathrm{a}}(x')$ où x' est le numéro de la case vers laquelle le joueur transite. On a donc le système

$$\begin{cases} m_{\mathrm{a}}(1) = 1 + (m_{\mathrm{a}}(2) + m_{\mathrm{a}}(4))/2 \\ m_{\mathrm{a}}(2) = m_{\mathrm{a}}(7) \\ m_{\mathrm{a}}(3) = m_{\mathrm{a}}(5) \\ m_{\mathrm{a}}(4) = 1 + (m_{\mathrm{a}}(5) + m_{\mathrm{a}}(6))/2 \\ m_{\mathrm{a}}(5) = 1 + (m_{\mathrm{a}}(6) + m_{\mathrm{a}}(7))/2 \\ m_{\mathrm{a}}(6) = m_{\mathrm{a}}(1) \\ m_{\mathrm{a}}(7) = 1 + (m_{\mathrm{a}}(8) + m_{\mathrm{a}}(9))/2 \\ m_{\mathrm{a}}(8) = m_{\mathrm{a}}(4) \\ m_{\mathrm{a}}(9) = 0 \end{cases} \qquad \text{de solution} \qquad \begin{pmatrix} m_{\mathrm{a}}(1) \\ m_{\mathrm{a}}(2) \\ m_{\mathrm{a}}(3) \\ m_{\mathrm{a}}(4) \\ m_{\mathrm{a}}(5) \\ m_{\mathrm{a}}(6) \\ m_{\mathrm{a}}(7) \\ m_{\mathrm{a}}(8) \\ m_{\mathrm{a}}(9) \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \\ 7 \\ 8 \\ 7 \\ 7 \\ 5 \\ 8 \\ 0 \end{pmatrix}.$$

Le nombre de tours moyens pour achever le jeu est donc 7.

Reprenons cette question avec ce second diagramme, qui est markovien,



où sont identifiées les cases jointes de manière instantanée. L'espace d'états est

$$E = \big\{\{1,6\}, \{3,5\}, \{4,8\}, \{2,7\}, 9\big\},$$

et $m_{\mathrm{a}}(x) = m_{\mathrm{a}}^9(x) = \mathbb{E}^x[T^9]$, $x \in E$, vérifient $m_{\mathrm{a}}(9) = 0$ et

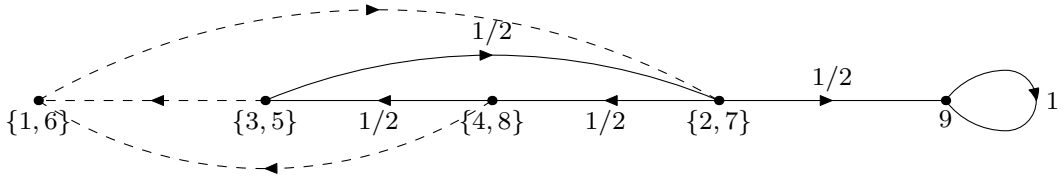
$$m_{\mathrm{a}}(x) = 1 + \sum_{y \in E, y \neq 9} P(x,y) m_{\mathrm{a}}(y),$$

soit

$$\begin{cases} m_{\rm a}(\{1,6\}) = 1 + \big(m_{\rm a}(\{3,5\}) + m_{\rm a}(\{2,7\})\big)/2 \\ m_{\rm a}(\{3,5\}) = 1 + \big(m_{\rm a}(\{1,6\}) + m_{\rm a}(\{2,7\})\big)/2 \\ m_{\rm a}(\{4,8\}) = 1 + \big(m_{\rm a}(\{1,6\}) + m_{\rm a}(\{3,5\})\big)/2 \\ m_{\rm a}(\{2,7\}) = 1 + m_{\rm a}(\{4,8\})/2 \end{cases} \qquad \text{de solution} \qquad \begin{pmatrix} m_{\rm a}(\{1,6\}) \\ m_{\rm a}(\{3,5\}) \\ m_{\rm a}(\{4,8\}) \\ m_{\rm a}(\{2,7\}) \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \\ 8 \\ 5 \end{pmatrix}$$

Nous retrouvons ainsi le même résultat avec, cette fois-ci, le formalisme markovien standard.

(ii) Nous partons de 5 sur le plateau, ou de $\{3,5\} \in E$. Pour ne pas retourner en 1, ou en $\{1,6\}$, avant d'atteindre 9, les seules transitions autorisées sont celles qui sont indiquées en traits pleins sur la figure suivante :



On atteint 9 en, ou bien 2 pas, ou bien 6 pas, ou bien 9 pas, ..., ou bien $2+3n$ pas, etc. L'événement considéré s'identifie à la réunion de tels parcours et sa probabilité est donc la somme des probabilités correspondantes :

$$\mathbb{P}^{\{3,5\}}\big\{T^9 < T^{\{1,6\}}\big\} = \sum_{n\geqslant 0} \Big(\frac{1}{2}\Big)^{2+3n} = \frac{1}{4}\sum_{n\geqslant 0}\Big(\frac{1}{8}\Big)^n = \frac{1}{4}\times\frac{1}{1-1/8} = \frac{1}{4}\times\frac{8}{7} = \frac{2}{7}\,.$$

Nous aurions aussi pu utiliser plus explicitement la propriété de Markov en posant $p_i = \mathbb{E}^i\{T^9 < T^1\}$, $i \in E$. On a clairement $p_{\{1,6\}} = 0$, $p_9 = 1$, et $p_i = \sum_{j\in E} P(i,j)p_j$, pour i différent de $\{1,6\}$ (9 est absorbant et ne pose pas de problème). On obtient alors le système

$$\begin{cases} p_{\{3,5\}} = 1/2 \times 0 + 1/2 \times p_{\{2,7\}} \\ p_{\{4,8\}} = 1/2 \times 0 + 1/2 \times p_{\{3,5\}} \\ p_{\{2,7\}} = 1/2 \times 1 + 1/2 \times p_{\{4,8\}} \end{cases} \qquad \text{de solution} \qquad \begin{pmatrix} p_{\{3,5\}} \\ p_{\{4,8\}} \\ p_{\{2,7\}} \end{pmatrix} = \begin{pmatrix} 2/7 \\ 1/7 \\ 4/7 \end{pmatrix}$$

Exercice 6 (ruine du joueur). — On considère une chaîne de Markov $(X_n)_{n\geqslant 0}$ d'espace d'états $E = \mathbb{N}$ et de matrice de transition P vérifiant

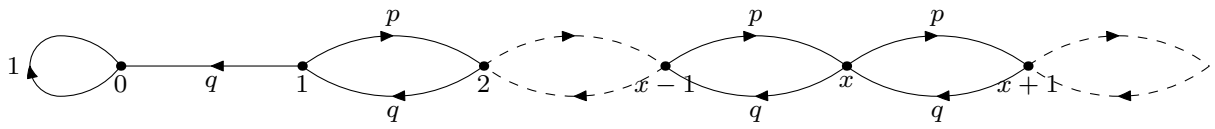
$$P(0,0) = 1, \qquad \text{et} \qquad P(x,x+1) = p, \quad P(x,x-1) = 1-p = q, \quad \text{pour tout } x \in \{1,2,\ldots\},$$

les autres coefficients étant nuls, $p \in [0,1]$ étant un réel donné. La variable X_n représente la fortune d'un joueur à l'instant n jouant contre un adversaire infiniment riche, la probabilité $p_{\rm a}^0(x) = \mathbb{P}^x\{T^0 < \infty\}$ est alors la probabilité de perdre avec une fortune initiale $x \in \mathbb{N}$. On note $k^0(x) = \mathbb{E}^x[T^0]$.

(i) Montrer que si $p \leqslant q$, $p_{\rm a}(x) = 1$ pour tout $x \in \mathbb{N}$ (distinguer les cas $p < q$ et $p = q = 1/2$), et si $p > q$, $p_{\rm a}^0(x) = (q/p)^x$. On établira une relation de récurrence entre les $p_{\rm a}^0(x)$.

(ii) Montrer que si $p < q$ on a $k^0(x) = x/(1-2p)$, et si $p \geqslant q$ on a $k^0(x) = \infty$ (distinguer les cas $p > q$ et $p = q = 1/2$). On établira une relation de récurrence entre les $k^0(x)$.

Correction. — Représentons le diagramme de transition



(i) Soit $p_{\mathrm{a}}^0(x) = \mathbb{P}^x\{\text{atteindre } 0\} = \mathbb{P}^x\{T^0 < \infty\}$. Nous savons que c'est la solution positive minimale du système

$$p_{\mathrm{a}}(0) = 1, \qquad p_{\mathrm{a}}(x) = p \times p_{\mathrm{a}}(x+1) + q \times p_{\mathrm{a}}(x-1), \quad x \geqslant 1.$$

Ce système définit une suite récurrente linéaire d'équation caractéristique

$$pr^2 - r + q = 0 \qquad (\Delta = 1 - 4pq \geqslant 0)$$

qui admet pour racines 1 (évident) et q/p (s'en déduit immédiatement). Lorsque $q/p \neq 1$, la suite est alors de la forme

$$p_{\mathrm{a}}(x) = a + b\Big(\frac{q}{p}\Big)^x$$

avec $p_{\mathrm{a}}(0) = a + b = 1$.

- Si $q > p$, comme h est bornée, on a $b = 0$ et donc $p_{\mathrm{a}}^0(x) = a = p_{\mathrm{a}}^0(0) = 1$: la ruine est sûre.
- Si $q < p$, écrivons $p_{\mathrm{a}}(x) = (q/p)^x + a(1 - ((q/p)^x)$. La solution positive minimale est obtenue pour $a = 0$, et donc $p_{\mathrm{a}}^0(x) = (q/p)^x$, notamment la ruine n'est pas certaine.
- Si $p = q = 1/2$, la solution s'écrit $p_{\mathrm{a}}^0(x) = a + bx$. Puisqu'elle doit être bornée, on a $b = 0$ et finalement $p_{\mathrm{a}}^0(x) = a = p_{\mathrm{a}}^0(0) = 1$: la ruine est certaine.

(ii) D'après le théorème précédent, k^0 est la solution positive minimale du système

$$k(0) = 0, \qquad k(x) = 1 + p \times k(x+1) + q \times k(x-1), \qquad x \geqslant 1.$$

Nous connaissons déjà les solutions du système homogène

$$p \times k(x+1) - k(x) + q \times k(x-1) = 0, \qquad x \geqslant 1.$$

Cherchons donc une solution particulière de la forme $k_*(x) = \alpha x$. On obtient alors $\alpha x = 1 + p\alpha x + p\alpha + q\alpha x - q\alpha$ soit $0 = 1 + \alpha(p - q)$ (rappelons que $p + q = 1$), et alors $\alpha = 1/(q - p) = 1/(1 - 2p)$ si $p \neq q$ ou encore $p \neq 1/2$. Ainsi $k_*(x) = x/(1 - 2p)$ est une solution positive lorsque $p \neq 1/2$. Toute solution s'écrit dans ce cas

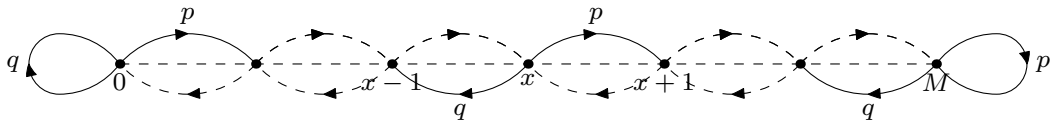
$$k(x) = a + b(q/p)^x + k_*(x) = a + b(q/p)^x + x/(1 - 2p) = b\big((q/p)^x - 1\big) + x/(1 - 2p).$$

puisque $k(0) = a + b = 0$.

- Si $p < q$, ou encore $p < 1/2$, les solutions sont du signe de b pour x assez grand. Pour qu'elles soient positives, il est donc nécessaire que $b \geqslant 0$, ce qui est aussi suffisant. La solution positive minimale est donc obtenue avec $b = 0$, soit $k^0(x) = x/(1 - 2p)$, $x \geqslant 1$.
- Si $p > q$, nous savons que $p_{\mathrm{a}}^0(x) = \mathbb{P}^x\{T^0 < \infty\} < 1$, par conséquent $\mathbb{P}^x\{T^0 = \infty\} > 0$, ce qui implique que $k^0(x) = \mathbb{E}^x[T^0] = \infty$ pour tout $x \geqslant 1$. Une autre façon de le voir est de constater que toute fonction finie $k(x) = b((q/p)^x - 1) + x/(1 - 2p)$ est du signe de $1/(1 - 2p) < 0$ pour x assez grand, et donc qu'il n'existe pas de solution positive finie. le minimum de l'ensemble vide étant l'infini, on retrouve la même conclusion.
- Si $p = q = 1/2$, alors on recherche une solution particulière de la forme $k_*(x) = \alpha x^2$, soit $\alpha x^2 = 1 + \alpha/2(x^2 + 2x + 1) + \alpha/2(x^2 - 2x + 1)$ ou encore $0 = 1 + \alpha$. Les solutions s'écrivent donc $k(x) = a + bx - x^2$, $x \geqslant 1$, avec $a = 0$ puisque $k(0) = 0$ et b quelconque. Aucune de ces solutions est positive puisqu'elles tendent toutes vers $-\infty$ en ∞. Nous avons donc $k^0(x) = \infty$ pour tout $x \geqslant 1$.

Exercice 7. — Soient l'espace d'états $E = \{0, 1, \ldots, M\}$, $p = 1 - q \in]0, 1[$, et la matrice de transition P définie par

$$P(0,0) = q, \qquad P(x+1, x) = p, \qquad 0 \leqslant i \leqslant M - 1\,;$$
$$P(x, x-1) = q, \quad 1 \leqslant x \leqslant M\,; \quad P(M, M) = p,$$



les autres coefficients étant nuls. Montrer que P est réversible par rapport à la mesure $\mu = ((p/q)^x)_{0\leqslant x\leqslant M}$. En déduire une mesure de probabilité par rapport à laquelle P soit réversible.

Correction. — Il est clair que la chaîne est irréductible. Comme l'espace d'états est fini, elle est (positivement) récurrente. Elle admet alors une unique mesure de probabilité invariante.

Soit $\mu = ((p/q)^x)_{0\leqslant x\leqslant M}$. Si $x \in \{1, \dots, M-1\}$, on ne peut transiter que vers $x-1$ et $x+1$. On a d'une part

$$\mu\{x\}P(x, x-1) = (p/q)^x \times q = p^x/q^{x-1} = (p/q)^{x-1} \times p = \mu\{x-1\}P(x-1, x),$$

et d'autre part

$$\mu\{x\}P(x, x+1) = (p/q)^x \times p = p^{x+1}/q^x = (p/q)^{x+1} \times q = \mu\{x+1\}P(x+1, x),$$

et ainsi, pour $x \in \{1, \dots, M-1\}$, $\mu\{x\}P(x, y) = \mu\{y\}P(y, x)$ pour tout $y \in E$. Considérons les bords : de 0 on peut transiter vers 0 ou vers 1 et on a

$$\mu\{0\}P(0, 1) = 1 \times p = p/q \times q = \mu\{1\}P(1, 0)$$

(et évidemment $\mu\{0\}P(0, 0) = \mu\{0\}P(0, 0)$) ; et de M on peut transiter vers M ou vers $M-1$ et on a

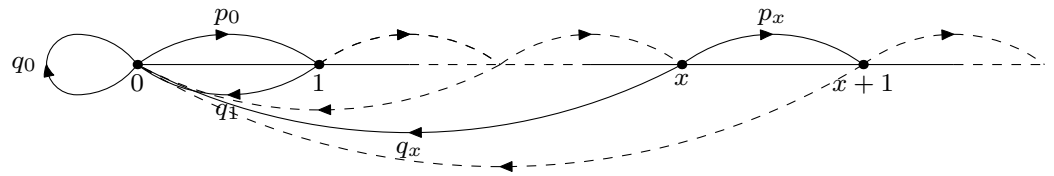
$$\mu\{M\}P(M, M-1) = (p/q)^M \times q = p^M/q^{M-1} = (p/q)^{M-1} \times p = \mu\{M-1\}P(M-1, M)$$

(et évidemment $\mu\{M\}P(M, M) = \mu\{M\}P(M, M)$). Ces deux dernières vérifications étaient finalement contenues dans celle qui précédait. On obtient que μ est une mesure positive, finie, par rapport à laquelle P est réversible, elle est donc invariante. Soit $\pi = \mu/\mu(E)$, plus précisément

$$\pi\{x\} = (p/q)^x \times \frac{1-2p}{q - q(p/q)^{M+1}}, \qquad x \in E = \{0, \dots, M\},$$

alors π est l'unique mesure de probabilité invariante de la chaîne.

Exercice 8. — Nous considérons la matrice de transition sur $\mathbb{N}$ représentée par la figure suivante :



c'est-à-dire que pour tout $x \in \mathbb{N}$, $P(x, x+1) = p_x$, $P(x, 0) = q_x = 1 - p_x$, $p_x \in [0, 1]$, les autres coefficients étant nuls. Sous l'hypothèse de la convergence du produit infini $\prod_{i=0}^{\infty} p_x$, montrer que cette matrice de transition n'a pas de mesure positive invariante (autre que nulle).

Correction. — Soit π une mesure sur $\mathbb{N}$. Au bout d'un pas, la mesure en 0 est $\pi P\{0\} = \sum_{x\in\mathbb{N}} \pi\{x\}P(x, 0) = \sum_{x\in\mathbb{N}} \pi\{x\}q_x$, alors qu'en $x \in \mathbb{N}^*$ la seule contribution ne peut venir que de $x-1$, $\pi P\{x\} = \pi\{x-1\}p_{x-1}$.

Ce que nous venons d'écrire se retrouve sur la formule générale $\pi P\{x\} = \sum_{y\in E} \pi\{y\}P(y, x)$, mais c'est bien aussi dans des cas simples de se rappeler comment marche une telle formule.

Si la mesure π est invariante, on a $\pi\{x\} = \pi P\{x\} = \pi\{x-1\}p_{x-1}$ pour tout $x \geqslant 1$, et ainsi

$$\pi\{x\} = \pi\{0\} \times p_0 \times \cdots \times p_{x-1} \qquad \text{pour tout } x \geqslant 1.$$

Toujours en supposant l'invariance, on a de plus $\pi\{0\} = \pi P\{0\}$, soit

$$\pi\{0\} = \sum_{x\in\mathbb{N}} \pi\{0\} \times p_0 \times \cdots \times p_{x-1} \times q_x = \pi\{0\} \times \sum_{x\in\mathbb{N}} p_0 \times \cdots \times p_{x-1} \times q_x$$

Mais, en examinant la dynamique de la chaîne, on constate que

$$\sum_{x\in\mathbb{N}} p_0 \times \cdots \times p_{x-1} \times q_x = \mathbb{P}^0\{T_{\mathrm{e}}^0 < \infty\}$$

car chaque terme de la somme est la probabilité d'une excursion hors de $\{0\}$ et la somme épuise tous les cas de figures. Or,

$$\mathbb{P}^0\{T_{\mathrm{e}}^0 = \infty\} = \lim_{x\to\infty} \mathbb{P}^0\{T_{\mathrm{e}}^0 \geqslant x\} = \lim_{x\to\infty} \prod_{y=0}^{x-1} p_y = \prod_{y=0}^{\infty} p_y = p$$

produit infini convergent par hypothèse (vers une valeur p strictement positive). Ainsi

$$\sum_{x\in\mathbb{N}} p_0 \times \cdots \times p_{x-1} \times q_x = \mathbb{P}^0\{T_{\mathrm{e}}^0 < \infty\} < 1$$

et comme $\pi\{0\} = \pi\{0\} \times \mathbb{P}^0\{T_{\mathrm{e}}^0 < \infty\}$, on a, ou bien $\pi\{0\} = 0$, ou bien $\pi\{0\} = \infty$. D'où la conclusion : il n'existe pas de mesure positive invariante non triviale.

Exercice 9. — Soit $(X_n)_{n\geqslant 0}$ une chaîne de Markov sur $E = \{0, 1, 2, \ldots\}$ dont les probabilités de transition sont données par

$$P(0,1) = 1, \qquad P(x,x+1)+P(x,x-1) = 1, \qquad P(x,x+1) = \left(\frac{x+1}{x}\right)^2 P(x,x-1), \quad x \geqslant 1.$$

Montrer que si $X_0 = 0$ alors la probabilité que $X_n \geqslant 1$ pour tout $n \geqslant 1$ est égale à $6/\pi^2$.

Correction. — Soit $X = (X_n)_{n\geqslant 0}$ une telle chaîne de Markov. On cherche la probabilité partant de x d'atteindre 0. Notons $p_x = P(x, x+1)$ et $q_x = 1 - p_x = P(x, x-1)$. Nous avons

$$p_{\mathrm{a}}^0(0) = 1, \qquad p_{\mathrm{a}}^0(x) = p_x p_{\mathrm{a}}^0(x+1) + q_x p_{\mathrm{a}}^0(x-1).$$

Pour passer d'une récurrence linéaire d'ordre 2 à une récurrence linéaire d'ordre 1, nous considérons la suite « dérivée » qui se traduit ici par les différences finies. Posons $u(x) = p_{\mathrm{a}}^0(x-1) - p_{\mathrm{a}}^0(x)$, $x \geqslant 1$, qui vérifie alors $p_x u(x+1) = q_x u(x)$, d'où

$$u(x+1) = \frac{q_x}{p_x} \times \frac{q_{x-1}}{p_{x-1}} \times \cdots \times \frac{q_1}{p_1} u(1) = \left(\frac{x}{x+1}\right)^2 \times \left(\frac{x-1}{x}\right)^2 \times \cdots \times \left(\frac{1}{2}\right)^2 u(1) = \left(\frac{1}{x+1}\right)^2 u(1)$$

Ainsi, par dominos (de manière imagée), on obtient

$$p_{\mathrm{a}}^0(0) - p_{\mathrm{a}}^0(x) = \sum_{y=1}^{x} u(y) = \sum_{y=1}^{x} \frac{1}{y^2} \times u(1).$$

ou encore

$$p_{\mathrm{a}}^0(x) = p_{\mathrm{a}}^0(0) - u(1) \sum_{y=1}^{x} \frac{1}{y^2} = 1 - u(1) \sum_{y=1}^{x} \frac{1}{y^2}$$

Le facteur $u(1)$ n'est pas encore identifié. Ce que nous savons est que p_{a}^0 est la solution minimale positive du système linéaire dont nous venons d'expliciter les solutions finies : $p_{\mathrm{a}}(x) = 1 - a\sum_{y=1}^{x} 1/y^2$, $x \geqslant 0$ (une somme vide étant nulle). Pour qu'une telle fonction h soit positive, nous devons avoir pour tout $x \geqslant 0$, $0 \leqslant 1 - a\sum_{y=1}^{x} 1/y^2$, soit $a \leqslant 1/\sum_{y=1}^{x} 1/y^2$ pour tout $x \geqslant 0$ et donc $a \leqslant 1/\sum_{y=1}^{\infty} 1/y^2 = 1/(\pi^2/6) = 6/\pi^2$; et réciproquement, si a vérifie

cette inégalité, la solution est positive. On constate que la solution positive minimale est obtenue pour $a = 6/\pi^2$. Ainsi, $p_{\mathrm{a}}^0(x) = 1 - 6/\pi^2 \sum_{y=1}^{x} 1/y^2$, $x \geqslant 0$. En particulier,

$$1 - 6/\pi^2 = p_{\mathrm{a}}^0(1) = \mathbb{P}^1\{T^0 < \infty\} = \mathbb{E}^0\{T_{\mathrm{e}}^0 < \infty\}$$

et $\mathbb{E}^0\{T_{\mathrm{e}}^0 = \infty\} = 6/\pi^2$, qui est bien la probabilité que X_n demeure supérieur ou égal à 1 pour tout $n \geqslant 1$ lorsque $X_0 = 0$.

Exercice 10. — Soit P une matrice de transition sur un espace d'états fini E.

(i) Montrer qu'une mesure de probabilité π est invariante si et seulement si

$$\pi \times \left(\mathrm{Id} - P + \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \right) = (1, \dots, 1).$$

(ii) En déduire que si P est irréductible, alors

$$M = \mathrm{Id} - P + \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$$

est inversible.

Remarque. — Cet exercice montre comment utiliser des méthodes matricielles pour déterminer la probabilité invariante d'une matrice de transition irréductible. Dans un cadre plus général, toujours avec E fini, il faudrait considérer chacune des sous-matrices associées à chaques classes récurrentes.

Correction. — (i) L'exercice est plus subtil qu'il n'y paraît. Il faut le prendre de manière purement algébrique au début. La matrice P est une matrice de transition. Nous noterons E' l'ensemble des vecteurs lignes et $\sigma : E' \to \mathbb{R}$ la forme linéaire qui à un vecteur associe la somme de ses coefficients, c'est-à-dire $\sigma(\gamma) = \gamma(1)$ où 1 est le vecteur colonne unaire ou la fonction constante égale à 1.

Pour tout vecteur ligne $\gamma \in E'$, on a évidemment $\sigma(\gamma\,\mathrm{Id}) = \sigma(\gamma)$, mais aussi $\sigma(\gamma P) = \sigma(\gamma)$ car P est une matrice de transition :

$$\sigma(\gamma P) = \sum_{x \in E} \gamma P\{x\} = \sum_{x \in E} \sum_{y \in E} \gamma\{y\} P(y, x) = \sum_{y \in E} \gamma\{y\} \sum_{x \in E} P(y, x) = \sum_{y \in E} \gamma\{y\} = \sigma(\gamma).$$

Ainsi, pour tout $\gamma \in E'$, $\sigma(\gamma(\mathrm{Id} - P)) = 0$.

Soit U la matrice unaire. Pour tout vecteur ligne γ, γU est le vecteur ligne dont tous les coefficients sont égaux à la somme des coefficients de γ, c'est-à-dire

$$\gamma(U) = (\sigma(\gamma), \dots, \sigma(\gamma)) = \sigma(\gamma) \times (1, \dots, 1) = \sigma(\gamma) \times 1'$$

où $1'$ est le vecteur ligne unaire. En passant, on a $\sigma(\gamma U) = n \times \sigma(U)$.

Finalement, pour tout $\gamma \in E'$, $\sigma(\gamma(\mathrm{Id} - P + U)) = n \times \sigma(\gamma)$.

Un vecteur ligne $\gamma \in E'$ est invariant si et seulement si $\gamma P = \gamma$, soit $\gamma(\mathrm{Id} - P) = 0'$ le vecteur ligne nul, ou encore, si et seulement si $\gamma(\mathrm{Id} - P + U) = \gamma U = \sigma(\gamma) \times 1'$.

Ainsi un vecteur ligne $\gamma \in E'$ est invariant et $\sigma(\gamma) = 1$ si et seulement si $\gamma(\mathrm{Id} - P + U) = 1'$. Ceci répond à la première question et prépare la réponse à la seconde.

(ii) Nous supposons la matrice de transition P irréductible. Comme l'espace d'états E est fini, nous savons qu'il existe une unique mesure de probabilité invariante $\pi \in E'$ et que tous ses coefficients sont strictement positifs.

Montrons que $M = (\mathrm{Id} - P + U)$ est inversible. Soit $\gamma \in E'$ tel que $\gamma M = 0'$ le vecteur ligne nul. Nous avons notamment $n \times \sigma(\gamma) = \sigma(\gamma M) = \sigma(0') = 0$, et donc $\sigma(\gamma) = 0$. Les coefficients

de γ étant finis et ceux de π étant strictement positifs, soit $\varepsilon > 0$ tel que $\pi_\varepsilon = \pi + \varepsilon\gamma \in E'$ a tous ses coefficients positifs ou nuls (prendre $0 < \varepsilon \leqslant \inf\{|\pi\{x\}/\gamma\{x\}| : x \in E\}$, infimum qui est strictement positif, voire infini). (Notons au passage que $\sigma(\pi_\varepsilon) = \sigma(\pi) + \varepsilon \times \sigma(\gamma) = 1 + \varepsilon \times 0 = 1$, et donc que π_ε est une mesure de probabilité sur E.) On a $\pi_\varepsilon M = \pi M + \varepsilon \times \gamma M = 1' + \varepsilon \times 0' = 1'$, ce qui montre que π_ε est invariante et $\sigma(\pi_\varepsilon) = 1$ (ce que nous avions déjà remarqué). Par unicité de la mesure de probabilité invariante $\pi_\varepsilon = \pi$, et donc $\varepsilon \times \gamma = 0'$, soit $\gamma = 0'$.

En conclusion, si $\gamma \in E'$ est tel que $\gamma M = 0'$, alors $\gamma = 0'$. Dans le langage courant de l'algèbre linéaire, ceci signifie que M' la matrice transposée de M est injective. Puisque nous sommes en dimension finie, M' est alors bijective et il en est de même de M (en utilisant l'invariance du déterminant par transposition par exemple).

Exercice 11. — Montrer qu'un état $x \in E$ est apériodique si et seulement si

$$n_x = \mathrm{pgcd}\big\{n \geqslant 0 : P^n(x,x) > 0\big\} = 1,$$

où pgcd est l'opérateur « plus grand commun diviseur ». (Lorsque n_x est en entier supérieur ou égal à 2, l'état x est dit périodique de période n_x.)

Correction. — Non disponible.

Exercice 12 (théorème de renouvellement). — Soit $(Y_n)_{n\geqslant 1}$ une suite de variables aléatoires à valeurs dans $\mathbb{N}^* = \{1, 2, \ldots\}$ indépendantes et de même loi. On suppose que

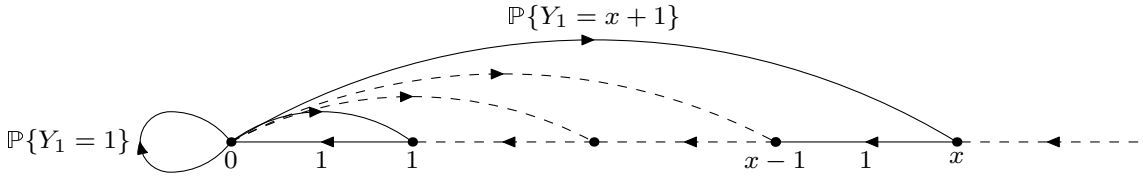
$$\mathrm{pgcd}\big\{n \geqslant 1 : \mathbb{P}\{Y_1 = n\} > 0\big\} = 1,$$

et on pose $\mu = \mathbb{E}[Y_1] \in \mathbb{R}_+ \cup \{\infty\}$.

Montrer que le processus $(X_n)_{n\geqslant 0}$ défini par $X_0 = 0$ et

$$X_n = \inf\{m \geqslant n : \exists\, k \geqslant 0,\, m = Y_1 + \cdots + Y_k\} - n, \qquad n \geqslant 1,$$

est une chaîne de Markov de matrice de transition P vérifiant $P(0,x) = \mathbb{P}\{Y_1 = x+1\}$, $x \geqslant 0$, et $P(x, x-1) = 1$, $i \geqslant 1$, les autres coefficients étant nuls.



On pourra poser $S_0 = 0$ et $S_k = Y_1 + \cdots + Y_k$ pour $k \geqslant 1$, et utiliser le fait que, si $S_{k-1} < n \leqslant S_k$, alors $X_n = S_k - n = Y_k - (n - S_{k-1})$.

Déterminer $\lim_{n\to\infty} \mathbb{P}\{X_n = 0\}$ et en déduire que

$$\mathbb{P}\{\exists\, k \geqslant 0 : n = Y_1 + \cdots + Y_k\} \longrightarrow \frac{1}{\mu} \qquad \text{quand } n \text{ tend vers l'infini.}$$

Correction. — Non disponible.

Il est clair que chaque X_n est une variable aléatoire à valeurs dans $\mathbb{N}$. Soient $n \in \mathbb{N}$ et $x_1, \ldots, x_n \in \mathbb{N}$. On a

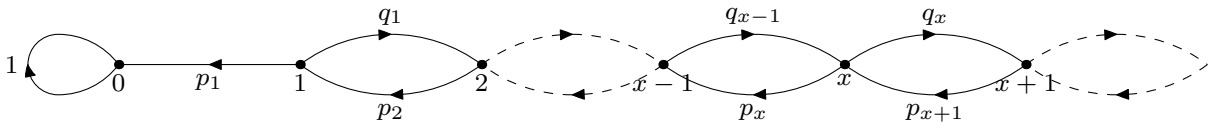
$$\mathbb{P}\{X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n\} = \mathbb{P}\{X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n\}$$

Exercice 13 (naissance et mort). — On considère $E = \mathbb{N}$, $(p_x)_{x\geqslant 1}$ une suite de réels dans $[0,1]$ et la matrice de transition P définie par

$$P(0,0) = 1, \qquad \text{et} \qquad P(x, x-1) = p_x, \quad P(x, x+1) = 1 - p_x = q_x, \quad \text{pour tout } x \geqslant 1,$$

tous les autres coefficients étant nuls. Chaque variable X_n représente la taille de la population après n changements, si la population est de taille x, p_x est la probabilité qu'il y ait une mort,

$q_x = 1 - p_x$ est celle qu'il y ait une naissance. La probabilité d'extinction de la population est alors $p_{\mathrm{a}}(x) = p_{\mathrm{a}}^0(x)$ lorsque la taille de la population est initialement x.



En posant $u(x) = p_{\mathrm{a}}(x-1) - p_{\mathrm{a}}(x)$, montrer que $p_x \times u(x+1) = q_x \times u(x)$, puis que

$$u(x+1) = \frac{q_x}{p_x}\, u(x) = \frac{q_x \times \cdots \times q_1}{p_x \times \cdots \times p_1}\, u(x) = \gamma(x) \times u(1).$$

En constatant que $u(1) + \cdots + u(x) = p_{\mathrm{a}}(0) - p_{\mathrm{a}}(x)$, en déduire que

$$p_{\mathrm{a}}(x) = 1 - u(1)(\gamma(0) + \cdots + \gamma(x-1))$$

où $\gamma(0) = 1$. Il reste à déterminer $u(1)$.

Si $\sum_{x=0}^{\infty} \gamma(x) = \infty$, montrer que $u_1 = 0$ et qu'ainsi $p_{\mathrm{a}}(x) = 1$ pour tout x.

Si $\sum_{x=0}^{\infty} \gamma(x) < \infty$, montrer que $u_1 = 1/\sum_{x=0}^{\infty} \gamma(x)$ est le choix qui donne les $\gamma(x)$ et que

$$p_{\mathrm{a}}(x) = \frac{\sum_{y=x}^{\infty} \gamma(y)}{\sum_{y=0}^{\infty} \gamma(y)}.$$

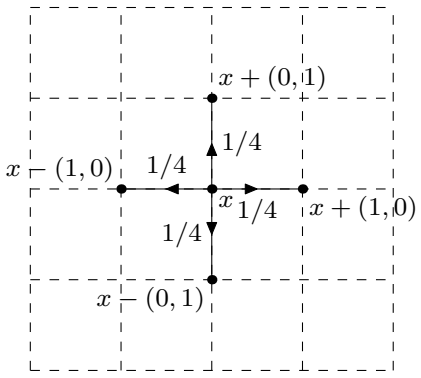
Ainsi, lorsque $x \geqslant 1$, on a $p_{\mathrm{a}}(x) < 1$, donc la population survit avec une probabilité strictement positive.

Exercice 14 (marche aléatoire simple dans $\mathbb{Z}$, suite). — Montrer que si $1 > p > q > 0$, alors l'unique classe communicante est $\mathbb{Z}$ et elle est transiente. Montrer que si $p = q = 1/2$, l'unique classe communicante est $\mathbb{Z}$ et qu'elle est récurrente. On pourra dans ce dernier cas utiliser la formule de Stirling

$$m\,! \approx \sqrt{2\pi m} \times \Big(\frac{m}{\mathrm{e}}\Big)^m \qquad \text{lorsque } m \to \infty.$$

Correction. — Non disponible.

Exercice 15 (marche aléatoire simple symétrique dans $\mathbb{Z}^2$). — La marche aléatoire symétrique dans $\mathbb{Z}^2$ est une chaîne de Markov telle que, pour tout n, X_n a probabilité $1/4$ de passer à l'un des 4 sites voisins



autrement dit, sa matrice de transition P est définie, pour tout $x \in \mathbb{Z}^2$, par

$$P\big(x, x+(1,0)\big) = P\big(x, x+(0,1)\big) = P\big(x, x-(1,0)\big) = P\big(x, x-(0,1)\big) = \frac{1}{4}\,,$$

les autres coefficients étant nuls.

On pose X_n^+ la projection orthogonale de X_n sur la diagonale $\{x = y\}$ et X_n^- la projection orthogonale de X_n sur l'anti-diagonale $\{x = -y\}$. Montrer que X_n^+ et X_n^- sont des marches aléatoires simples symétriques dans $\frac{1}{\sqrt{2}}\mathbb{Z}$, qui sont de plus indépendantes. En déduire que

$$P^{2n}(0,0) = \left(\binom{2n}{n} \times \left(\frac{1}{2}\right)^{2n}\right)^2 \approx \frac{1}{\pi n} \qquad \text{quand } n \to \infty,$$

puis que $(X_n)_{n\geqslant 0}$ est récurrente, c'est à-dire qu'il n'y a qu'une seule classe communicante (irréductibilité), et que celle-ci est récurrente.

Correction. — Non disponible.

Exercice 16 (marche aléatoire simple symétrique dans $\mathbb{Z}^3$). — La marche aléatoire symétrique dans $\mathbb{Z}^3$ est une chaîne de Markov telle que, pour tout n, X_n a probabilité $1/6$ de passer à l'un des 6 sites voisins, autrement dit, sa matrice de transition P est définie par, pour tout $x \in \mathbb{Z}^3$,

$$\begin{aligned} P\big(x, x+(1,0,0)\big) = P\big(x, x+(0,1,0)\big) = P\big(x, x+(0,0,1)\big) \\ = P\big(x, x-(1,0,0)\big) = P\big(x, x-(0,1,0)\big) = P\big(x, x-(0,0,1)\big) = \frac{1}{6}\,, \end{aligned}$$

les autres coefficients étant nuls.

Montrer que

$$P^{2n}(0,0) = \binom{2n}{n}\left(\frac{1}{2}\right)^{2n} \times \sum_{\substack{i,y,k\geqslant 0\\ i+y+k=n}} \binom{n}{i\ y\ k}^2 \left(\frac{1}{3}\right)^{2n}$$

où

$$\binom{n}{i\ y\ k} = \frac{n\,!}{i\,!\,y\,!\,k\,!}$$

est un coefficient multinomial. En utilisant la formule de Stirling, montrer que

$$\sum_{n\geqslant 0} P^n(0,0) = \sum_{n\geqslant 0} P^{2n}(0,0) < \infty.$$

En déduire que $(X_n)_{n\geqslant 0}$ est transiente, c'est-à-dire qu'il n'y a qu'une seule classe communicante (irréductibilité), et que celle-ci est transiente.

Correction. — Non disponible.

Exercice 17 (décomposition du premier retour). — Soit T_{r}^y le premier temps de retour dans l'état y, et

$$f^{(n)}(x,y) = \mathbb{P}^x\{T_{\mathrm{r}}^y = n\}.$$

Montrer que

$$P^n(x,y) = \sum_{k=1}^{n} f^{(k)}(x,y) \times P^{n-k}(y,y) \qquad \text{pour tout } n \geqslant 1.$$

En déduire que

$$P(x,y)(s) = \delta(x,y) + F(x,y)(s) \times P(y,y)(s)$$

où

$$P(x,y)(s) = \sum_{n=0}^{\infty} P^n(x,y) \times s^n \qquad \text{et} \qquad F(x,y)(s) = \sum_{n=0}^{\infty} f^{(n)}(x,y) \times s^n.$$

En déduire que $\mathbb{P}^x\{T_{\mathrm{r}}^x < \infty\} = 1$ si et seulement si $\sum_{n=0}^{\infty} P_{i,i}^{(n)} = \infty$ et ce, sans utiliser le théorème correspondant.

Correction. — Non disponible.